Demand and Decision

Three Empirical Essays on Insurance Demand
and the Transition into Self-Employment

Inaugural-Dissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2008

vorgelegt von
Veronika Regina Scharfenberg

Referent:	Prof. Dr. Joachim Winter
Korreferent:	Prof. Dietmar Harhoff, Ph.D.
Promotionsabschlussberatung:	4. Februar 2009

Für Oliver

Acknowledgements

First of all, I would like to thank my supervisor, Joachim Winter. His encouraging support and guidance in technical matters as well as his friendship were so important for this thesis.
I am also very thankful to Dietmar Harhoff for agreeing to become the co-supervisor.

Besides, I am very grateful to my colleagues from the Seminar für Empirische Wirtschaftsforschung for their helpful comments and especially to Helmut Farbmacher who helped me a lot collecting the data set for the first chapter.
I also thank Karl-Heinz Steiner from Allianz Private Krankenversicherungs-Aktiengesellschaft for providing a lot of useful background information regarding supplementary health insurance.

The Mikrozensus based chapters would not have been possible without the support of "Statistisches Bundesamt" in Wiesbaden who kindly provided the panel data sets right after they were published.

I gratefully acknowledge the financial support from Deutsche Forschungsgemeinschaft (DFG) from 2005 to 2008.

Finally, I would like to thank my parents and my husband Oliver for their support, especially in the harder, less motivated phases of my dissertation. You never ceased being positive and – most importantly – always believed in me finishing this project.

Table of Contents

List of Tables

List of Figures

Preface

"The strongest principle of growth lies in human choice."
- George Eliot[1] -

In Germany, neither economic growth nor the increase in population is sufficient to maintain social security systems as we know them. On the one hand this is why parts of this system are being privatized, e.g. the market for private supplementary long-term care (SLTC) insurance. On the other hand this lack of growth is the most important reason behind governmental subsidy programs to incentivize the population to create their own businesses. Both processes are subject to individual decisions: Do people respond to the incentives they are given? Do they enrol in new privatized markets? Why is demand in the market for private SLTC insurance so small?

Demand and Decision – this dissertation consists of three self-contained chapters that treat empirical studies about demand and individual decisions. They all have their own introduction and references and can be read independently of each other.

The three studies cover the subjects of demand and decision in several ways: The first two studies treat the same market and cover the same demand issues while the second and the third study are based on the Mikrozensus Panel, both analyzing individual decisions.

Chapters one and two both treat the German market for supplementary long-term care insurance: Since 1995, everybody in Germany has automatically been insured in compulsory long-term care (LTC) insurance. This insurance has been designed to cover basic risk – not as full cost coverage. Even though supplementary LTC (SLTC) insurance offers additional coverage for the LTC risk on an individual basis only few have an SLTC policy. Therefore the first two chapters concentrate on the factors which cause this low demand – and on the individual decisions of the customers behind demand.

The first chapter asks how products and suppliers in the rising market for private supplementary LTC insurance in Germany can be characterized. Are there any imperfections in competition? Should the government intervene? In order to answer these questions I collected a data set of German insurers and the contracts they offer. Based on this data set I run a hedonic regression on the insurance offers with the insurance price as the dependent variable and several

[1] Eliot (1876): p. 422 (see reference section of chapter 1).

independent variables such as individual characteristics like gender and age and contract-specific factors like standard waiting times and payments for informal care.
Customer characteristics and contract options account for 94% of the price variation in contracts. Premium differences are very high as well as search costs. As the market is far from perfect competition, a governmental objective should be to increase demand and to enhance competition. Therefore, the government should inform more about the limited cost coverage in compulsory LTC insurance, about individual possibilities to increase coverage and about the role of the SLTC insurance in pending reforms of the LTC insurance market.

While the first part concentrated on a comparison of insurance contracts in the German market, the second chapter starts with the individual decision of an insurance customer. Who enrols in the market? Answers are given based on a Probit regression for insurance demand. While age, sex, income and family characteristics are found to influence the decision whether to buy an SLTC insurance policy, health variables can not be shown to have a significant influence on the decision. In the main part of the chapter I explore supply side failures as a possible reason for the limited size of the market. Most market imperfections result in elevated prices[2] and / or "quantity rationing", i.e. contracts that do not cover the whole risk. However, I can rule out quantity rationing because in Germany, customers can freely choose the extent of coverage in their SLTC contract and quantity rationing is just not possible. Therefore, concentrating on insurance prices, I compare the expected present discounted values of lifelong premium and benefit payments to analyze premiums in the market – and to find out whether prices are indeed elevated. In order to compute the expected values in the model I estimate probabilities for one-year transitions into and out of LTC. After running an Ordered Probit regression on data from the Mikrozensus Panel 1996-1999 I take the predicted probabilities between the five states – health, care levels one, two and three, and death – to compute transition probabilities from any period to any period. In order to do so, I use Markov chains. The resulting probabilities are consistent with existing studies about the LTC risk, mortality rates and their differences for both sexes. Finally, I calculate the expected discounted values by using actual benefit payments and premiums from the insurance offers I collected for the study in the first chapter.
Premiums are shown to be very elevated, even when various forms of insurer costs are taken into consideration. These high prices can have their origin in supply side failures in the market, especially asymmetric information, imperfect competition and high transaction costs.

[2] See Brown and Finkelstein (2007) in the reference section of chapter 2.

The main link between the second and the third chapter is the recently published Mikrozensus Panel. Being based on the next version, the Mikrozensus Panel 2001-2004, this third study concentrates on decision, analyzing the factors which influence the individual decision to become self-employed.

As all new enterprises can help create jobs and consolidate social security systems new firm formation is the basis for growth and competitiveness of an economy. This is why the German government created a number of subsidies to support future entrepreneurs, one of them being the "Existenzgruendungszuschuss" in 2003, a fixed monthly amount designed to help unemployed persons to switch into self-employment.

The sizeable amount of public spending on support systems for new firms creates a large interest in identifying the factors which influence the individual decision to become self-employed. The reason is that beside the possibility to evaluate existing political instruments this analysis can also help design incentives for future new entrepreneurs.

This third chapter has two goals: after analysing the influence of microeconomic variables on the individual decision, I hold this influence constant and assess the effect of the introduction of the "Existenzgruendungszuschuss" in 2003. In order to account for unobserved individual heterogeneity in the analysis of the influencing factors, I estimate a Chamberlain's Random Effects Probit model. In this approach the correlation of the individual effect with the regressors is represented by adding the individual means of the variables to the equation.

Generally, male, unemployed or inactive persons are shown to be more likely to become self-employed than others. The decision is not significantly affected by personal wealth. Besides, a high household income has a positive effect on the probability of becoming self-employed while a high personal income decreases it. Concerning the influence of the EGZ, unemployed persons are – ceteris paribus – significantly more likely to become self-employed in 2003 than in the years before: governmental incentives seem to work.

Still, KfW (2008) showed that the businesses of previously unemployed persons tend to be less successful than those of the employed.[3] This is why the government should encourage especially employed and well educated persons to enter self-employment – and reconsider the economic contribution of previously unemployed new entrepreneurs.

[3] See KfW (2008) in the reference section of chapter 3.

1 Private Supplementary Long-Term Care Insurance in Germany: Characterization of a rising Market

1.1 Introduction

Since 1995, everybody in Germany is automatically insured in compulsory long-term care (LTC) insurance. To cover the costs, employees have to pay 0.85% of their income plus 0.25% if they are childless. The insurance has been designed to cover basic risk – not full cost coverage.[4] Social LTC insurance is mandatory for every German citizen holding public health insurance. Persons holding private health insurance are obliged to insure their LTC risk with a private LTC insurance policy.

According to social insurance law a nursing case is defined as a person who, caused by physical, intellectual or mental disease or disability, needs a substantial provision of nursing care and assistance with her activities of daily living (ADL's). Based on daily duration and type of assistance, e.g. nutrition, personal hygiene or mobility, three care levels and a hardship case are defined as shown in Table 1.1:

	Care Level 1	Care Level 2	Care Level 3	Hardship Case
Assistance	≥ 1.5 hrs / day	≥ 3 hrs / day	≥ 5 hrs / day	>> 5 hrs / day
Basic Care (as part of assistance time)	≥ 0.75 hrs / day	≥ 2 hrs / day	≥ 4 hrs / day	>> 4 hrs / day

Table 1.1: Care Level Definitions[5]

A person can only be assigned to a care level after she or her legal representative has applied for the LTC benefits. The official medical service of the German health insurances will then send an agent to the person's home to assess the amount of assistance the nursing case needs and to decide about her care level. Benefit payments depend on the care level and on the characteristics of the care giver – see Table 1.2 for details.

		Care Level			
Type of Care		**1**	**2**	**3**	**Hardship**
At Home (Choice between two options, combinations possible)	Professional Help	384	921	1,432	1,918
	Daily Allowance for informal helper	205	410	665	
Nursing Home	Professional Help	1,023	1,279	1,432	1,688

Table 1.2: LTC Payments in €[6]

Today these benefits cover about 53% of the costs people in nursing homes have to face. In 1999, it was 58%.[7] As benefits are not adapted to inflation the real coverage of risk has de-

[4] Bundesministerium für Gesundheit (2007b), p. 7.
[5] Bundesministerium für Gesundheit (2007b), p. 15.
[6] Bundesministerium für Gesundheit (2007a), p. 4.
[7] Statistisches Bundesamt (2007), p. 15; Bundesministerium für Gesundheit (2007a), p.5; Statistisches Bundesamt (2001), p. 13.

creased since the introduction of the Social LTC insurance. Due to the demographic structure the number of elder people in the country and thus the number of people in LTC will increase rapidly within the next 40 years.[8] As a consequence the funds available per nursing case will be reduced. Supplementary LTC (SLTC) insurance policies offer additional coverage for the LTC risk on an individual basis.

Insurance companies offer two types of SLTC insurance policies: the first type – "Pflegekostenversicherung" / care cost insurance – covers some of the actual costs that arise in LTC and that are not already covered by the compulsory LTC insurance. The second type – "Pflegetagegeldversicherung" / daily allowance insurance – defines a fixed amount of money as daily allowance which is paid in the case of LTC. This insurance is more flexible and thus more popular as benefits can also be used to pay relatives who take care of the insured person.

In this study I analyze the market for SLTC insurance. Comparing LTC institutionalization across several countries, I range the German system in an international context. To facilitate understanding of premium calculation, I give an overview on general health insurance calculation rules in Germany.
In the main part of the paper I characterize the German market for SLTC insurance. After a short presentation of recent numbers of the market I analyze the cost of SLTC insurance policies based on insurance offers I collected from private insurance companies in Germany.
Descriptive statistics show that the market is rather small with 800.000 insured persons. However, the market is growing fast with growth rates in benefit payments of up to 25% per year to 14.9 million € in 2005.
Prices for insurance policies rise with the age and so does the price variance in the market.

In order to compare premiums across contracts, I concentrate on daily allowance insurance policies. I run a hedonic regression on the insurance offers with the insurance price as the dependent variable and several independent variables such as individual characteristics like gender and age and contract-specific factors like standard waiting times and payments for informal care. This analysis explains a large portion of the observed variation in insurance premiums (r^2=94%). It indicates that premiums rise exponentially with the age and that women pay substantially more than men. Contracts that only pay benefits in care level three are shown to be less expensive than contracts with additional payments in other care levels.

[8] Schulz et al. (2004), p. 66.

A further hedonic regression includes dummy variables for the different companies that offer insurance policies in the market whereas the market leader is the reference company. Results show that most of the companies offer cheaper contracts than the market leader and that prices – even accounting for different contract characteristics – vary a lot in the market. The top ten health insurance companies in the market are shown to be as expensive as the rest of the market.

Since the market includes a total number of 26 companies there is clearly no monopolistic structure. Search costs are very high which shows – besides the complex price structure – that the market for private SLTC insurances is far from perfect competition. The main policy implication of the analysis is that the government should inform more about the limited cost coverage in compulsory LTC insurance, about individual possibilities to increase coverage and about the role of the SLTC insurance in pending reforms of the LTC insurance market.

There is large literature on private LTC insurance, mainly regarding the United States. Market size and market failure have been analyzed by Brown and Finkelstein (2004) who look at the supply side of the market or Finkelstein and McGarry (2006) who develop a test for asymmetric information in the American LTC market.

International comparison studies about LTC insurance markets include OECD (2005) for a general comparison between OECD countries and Colombo and Tapay (2004) who generally analyze private health insurance markets and focus on private SLTC insurance in some sections of their study.

In Germany, various publications describe mandatory LTC insurance, its problems and possible solutions. Schulz et al (2004) predict utilization rates of LTC and find a dramatic increase in the number of LTC cases. Häcker and Raffelhüschen (2004) point out that the current LTC system substantially lacks sustainability and that existing ideas to reform the system will not be able to solve the problem. However, no analysis of the German market for private SLTC insurance is known to the author.

1.2 International Comparison of Long-Term Care Systems

Most OECD members face the same demographic challenges but programmes for LTC coverage vary a lot across countries. Table 1.3 shows selected OECD countries and their LTC financing. Most countries offer a combination of in-kind benefits and cash payments for both home care and institutional care. In all countries except Korea, Japan and the US, LTC ser-

vices are available to all age groups. Mostly services are paid out of taxes; insurances as in the German system exist in some countries like Hungary or Japan.[9]

Country	Source of fund	Type of benefits
Australia	General Taxation	In-kind and Cash
Austria	General Taxation	Cash
Canada	General Taxation	In-kind
Germany	Insurance Contribution	In-kind and Cash
Hungary	General Taxation	In-kind and Cash
Ireland	Insurance Contribution	In-kind
Japan	Insurance Contribution and General Taxation	In-kind
Korea	General Taxation	In-kind
Luxembourg	Insurance Contribution	In-kind and Cash
Mexico	General Taxation	In-kind
Netherlands	Insurance Contribution	In-kind and Cash
New Zealand	General Taxation	In-kind
Norway	General Taxation	In-kind
Poland	General Taxation	In-kind and Cash
Spain	General Taxation	In-kind
Sweden	General Taxation	In-kind
Switzerland	Sickness / Old Age Insurance Funds and General Taxation	In-kind and Cash
United Kingdom	General Taxation	In-kind and Cash
United States	Insurance Contribution and General Taxation	In-kind

Table 1.3: Major public Programmes covering LTC in selected OECD Countries, 2003[10]

Country	Total Expenditure	Public Expenditure	Private Expenditure
Australia	1.19	0.86	0.33
Austria	n.a.	1.32	n.a.
Canada	1.23	0.99	0.24
Germany	1.35	0.95	0.4
Hungary	<0.30	<0.20	<0.1
Ireland	0.62	0.52	0.1
Japan	0.83	0.76	0.07
Korea	<0.30	<0.20	n.a.
Luxembourg	n.a.	0.52	n.a.
Mexico	<0.20	<0.10	<0.10
Netherlands	1.44	1.31	0.13
New Zealand	0.68	0.45	0.23
Norway	2.15	1.85	0.29
Poland	0.38	0.37	0
Spain	0.61	0.16	0.44
Sweden	2.89	2.74	0.14
Switzerland	1.54	n.a.	n.a.
United Kingdom	1.37	0.89	0.48
United States	1.29	0.74	0.54
Average	1.25	0.99	0.24

Table 1.4: Public and private Expenditure on LTC as a Percentage of GDP. 2000[11]

Total expenditure for LTC coverage also differs across countries. Expenses range from 0.2% to 3% of GDP – as seen in Table 1.4. In most countries, however, they are between 0.5% and 1.6% of GDP, Norway and Sweden being outliers at the top of the range. This table also shows that differences in funding principles do not always lead to differing expenditure outcomes.

[9] For a more detailed overview, see OECD (2005), p. 19-32.
[10] OECD (2005), p. 22-24.
[11] OECD (2005), p. 26.

Public expenditures are the most important source of LTC financing. However, compared to other ageing-related expenditures such as pensions or acute health care they are relatively low. In most countries private households have to make substantial co-payments or out-of-pocket spending for care. In the future, supplementary private insurance might thus play a stronger role in LTC coverage.

Nevertheless, as Colombo and Tapay (2004) point out, specialised private LTC insurance markets are absent or very limited in most OECD countries, even when no public coverage is offered. Besides Germany, markets for supplementary insurance have emerged in France, the US and the UK. High premiums have kept demand in the UK very low. French insurance policies offer fixed monthly benefits just as the German daily allowance insurance while in the US mostly occurring costs are reimbursed like in the German care cost insurance.

1.3 General Rules for calculating Health Insurance Premiums in Germany[12]

Private SLTC insurance is considered a part of the private health insurance sector. This is why the rules for the calculation of health insurance premiums also apply here. In the public sector insurance companies are obliged to contract possible customers whereas in the private sector they can choose whom to insure and whom to exclude. Private insurance companies conduct risk examinations before they allow a customer to sign a contract.

Premiums are calculated following a benefits principle: the expected present discounted value (EPDV) of all future benefits has to equal the EPDV of all future premiums. The insurer collects information about the customer's age, gender, current health status and sometimes additional risk factors like occupation and place of residence to determine the net EPDV of benefits. Future payments have to be discounted by a rate no higher than 3.5%. For the gross EPDV of benefits insurer costs and a security loading are added to the net amount. According to §7 KalV (Kalkulationsverordnung, German law for premium calculation), the security loading has to be higher than 5% of the gross premium. Furthermore, all additional charges for costs and security have to be calculated separately so that they cover the corresponding expected mean expenditures. Table 1.5 summarizes the premium calculation process.

The calculation of premiums "follows the type of assurance", i.e. premiums are calculated like assurance premiums; therefore the assurance's assumptions are transferred to the SLTC insurance: the current mortality rates and claims amounts per risk are valid throughout the whole duration of the contract. Furthermore, absence of inflation is assumed as well as cost-

[12] See Milbrodt (2005) for a more detailed description.

neutral medical improvement and constant insurer costs. Instead of the actually arising costs in the case of LTC, which are random, expected claims amounts per risk can be used for calculation. As a result of these assumptions, premiums can be held constant for the whole duration of the contract.

	Premium Composition	Premium Utilization
	Risk Premium	Benefits
+	Savings Premium	Premium Reserves
=	*Net Premium*	
+	Security Loading	Compensation of expenditures that exceed the actuarys' assumptions, usually between 5% and 10%
+	Insurer Costs	Acquisition costs, administrative costs, claims settlement costs
=	*Gross Premium*	

Table 1.5: Premium Composition[13]

Each cohort, i.e. each group of individuals of the same age and gender is paying for themselves only, so that there is no spread of risks between the elder and the younger. As benefits potentially rise with age, the cohorts have to accumulate aging reserves. In the first phase of the contract period, the build-up phase, they build savings that they can use in the second phase of the contract period, the withdrawal phase. See Figure 1.1 for an illustration.

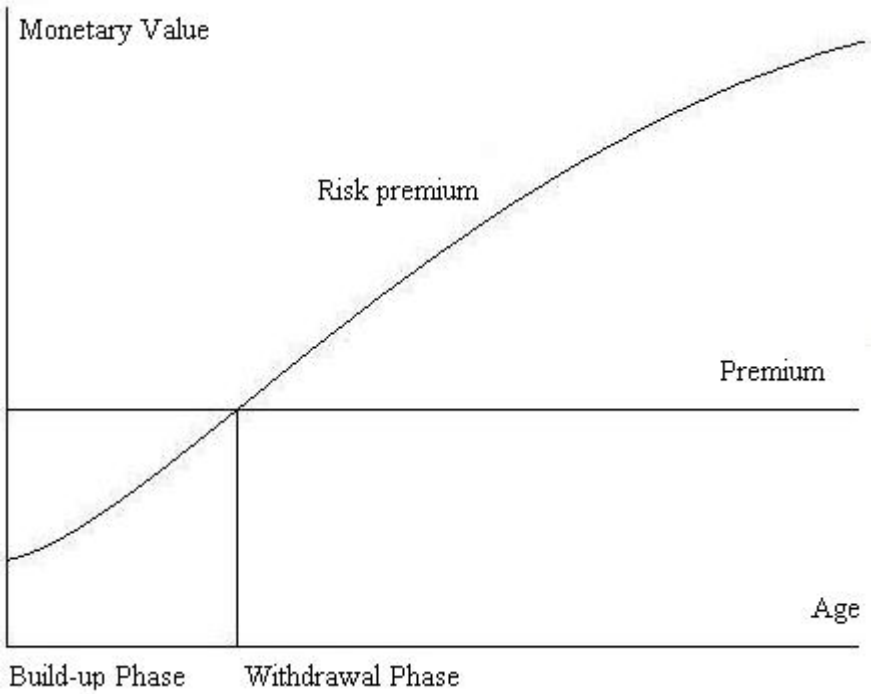

Figure 1.1: Two Phases of the Contract[14]

[13] According to Milbrodt (2005), p. 16.
[14] Milbrodt (2005), p.15.

As the assurance assumptions are not quite realistic insurance law allows for premium adjustments if circumstances change. In the SLTC insurance, such circumstances can be changing laws for compulsory LTC insurance or major changes in mortality tables.[15] Adjustments have to be confirmed by an independent actuary and they have to be based on premium adjustment clauses in the insurance contract.
Chapter 1.4 analyzes the currently available private SLTC insurance premiums in the German market based on customer characteristics and contract features as influencing factors.

1.4 The Market for Private Supplementary LTC Insurance

In Germany, we find different ways to insure the risk of becoming a nursing case in a health insurance company. Private insurance companies offer two kinds of insurance. The first type is called "Pflegekostenversicherung" – care cost insurance. It covers certain costs that arise in LTC. This insurance is not very flexible and often not recommended.[16] As it pays only as the need arises and depending on a lot of circumstances, it is very hard to model. Furthermore, only 21% of the persons in private SLTC insurance actually own this kind of insurance.[17]
This is why this study analyzes the second kind of insurance, the "Pflegetagegeldversicherung" – daily allowance insurance (SLTC insurance in this paper) – which pays a certain amount of money in the case of LTC. This insurance is very flexible as the benefits do not depend on the project the money is being used for. Thus, frail persons can pay money to their relatives who take care of them. This daily allowance insurance is better to model as the payment rules are alike for all insurance companies. Besides that this insurance is more popular in Germany (79%).

The SLTC insurance pays benefits irrespective of the payments of the compulsory LTC insurance and vice versa. There is no direct crowding out effect between the insurance plans as observed in the US by Brown et al. (2006). In order to avoid adverse selection, persons who buy an insurance policy have to be in good health. This is tested through a standard risk examination. In addition to that most of the contracts can only be closed if the applicant agrees to the condition that his family doctor may pass information about the applicant's health status to the contracting insurance company. If the applicant has severe health problems or diseases, insurance is either not possible or the premium is adapted to the elevated individual

[15] Email from Karl-Heinz Steiner (Allianz AG), 12-04-07.
[16] Finanztest (2006).
[17] Verband der privaten Krankenversicherung (PKV) (2007), p. 31.

risk. As no data is available about these individually negotiated premiums, these cases cannot be taken into account in this study.

1.4.1 The Data

Besides the official statistics – PKV (2000-2005) and Statistisches Bundesamt (2007) – I use data I collected between June and November 2006. The dataset includes information about the rates German insurers offer for SLTC insurance. Out of 53 private insurance companies in Germany at that time nine had regional offers only or were too small to be analyzed, 15 did not offer SLTC insurance and three – two of which do offer a daily allowance insurance – did not want to publish any data. The data used here is from the remaining 24 German insurers that offer daily allowance insurance. Table 1.6 shows those included insurance companies and their insurance plans.

Insurer	Plan 1	Plan 2	Plan 3
Allianz Private Krankenversicherungs-AG	PZT		
Alte Oldenburger Krankenversicherung	PT		
ARAG Krankenversicherungs-AG	Tarif 69		
Barmenia Krankenversicherung a.G.	PT1	PT3	
Bayerische Beamtenkrankenkasse	PflegeOPTIMAL	PflegeKOM-PAKT	
CENTRAL Krankenversicherung	EPTN1	EPTN2	
DBV-Winterthur Krankenversicherung	PTG 3	PTG DYN	
Debeka Krankenversicherungsverein	PVZ		
Deutscher Ring	PTG 1	PTG 2	
DEVK Krankenversicherungs-AG	PT/B	PT/B3	
DKV Deutsche Krankenversicherung	PET		
Gothaer Krankenversicherung	PT		
HanseMerkur Krankenversicherung	PTA		
HUK-Coburg-Krankenversicherung	PT	PT3	
INTER Krankenversicherung	PTN		
LVM Krankenversicherungs-AG	PZT		
Münchener Verein	Tarif 420	Tarif 423	
Nürnberger Krankenversicherung	PT		
PAX-Familienfürsorge Krankenversicherung	P EU		
R+V Krankenversicherung	PT	PT3	
SIGNAL Krankenversicherung	EPT		
Süddeutsche Krankenversicherung	PE1	PE2	PE3
uniVersa Krankenversicherung	PTK		
VICTORIA Krankenversicherung	PZ		

Table 1.6: Insurers and Insurance Plans

The dataset includes 24 insurers with, in total, 35 insurance plans. The price for an insurance policy basically depends on the customer's age and gender. Furthermore, contract-specific components influence the price, such as the benefits in each care level as shares of the benefit payment in care level three (e.g. 25% in level one, 50% in level two and 100% in level three), the maximum age at entry and the time after buying the insurance that the customer has to wait until he can receive benefits (usually up to 3 years).

Benefit payments are usually offered as multiples of a 5€ payment in care level three. There is a base price for a 5€ benefit unit. The total monthly premium is the combination of the base

price and the according factor to get the contracted benefit amount. E.g., if a 5€ daily allowance costs 0.50€ per month, a 20€ daily allowance costs 2€ per month. Premiums usually stay constant for a life-time[18] unless optional arrangements to increase coverage and thus premiums automatically after a certain period have been agreed on.

1.4.2 Descriptive Statistics of the Market for private supplementary LTC Insurance

The market for private SLTC insurance is still a small market compared to other private insurance markets.[19] In 2005 about 667.800 persons held daily allowance insurance, 8.8% more than in 2004.[20] The return from premiums was 170 Mio € when it was first recorded in 2005. Statistics before 2004 do not distinguish between daily allowance insurance and care cost insurance. For overall private SLTC insurance, the number of insured persons doubled from 380.000 in 1995 to 750.000 in 2003 and increased to 830.000 in 2005. Figure 1.2 illustrates this development.

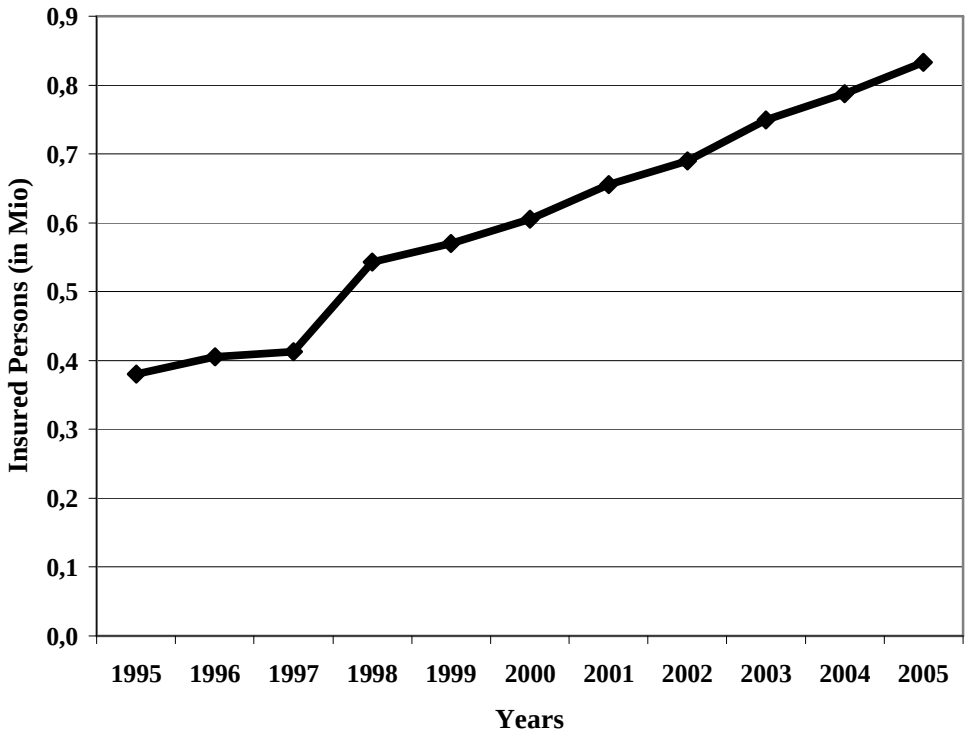

Figure 1.2: Insured Persons (in Mio)[21]

Growth rates in benefit payments are even higher as more and more persons reach LTC age. Information is available back to 7.8 million € in 2002 and shows a 20-25% growth rate up to

[18] See general assumptions in section 2.
[19] See Scharfenberg (2008) for a detailed discussion of the limited SLTC market size.
[20] See PKV (2006), p. 31, 36, 49, 52, 86 for the following statistics.
[21] PKV (2006), p. 86f.

14.9 million € in 2005. The raise in benefit payments of 22.7% from 11.9 Mio € in 2004 to 14.6 Mio € in 2005 was the highest increase in benefit payments in all private insurances in that period. Figure 3 shows the benefit development over the years 2002-2005.

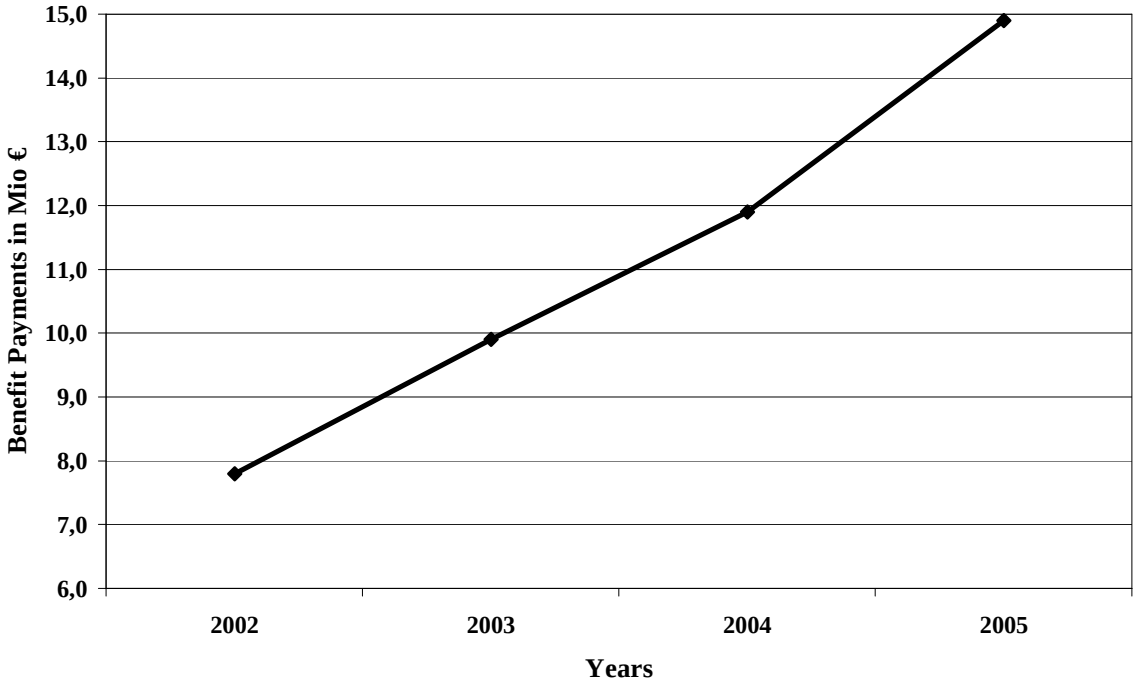

Figure 1.3: Benefit Payments in Mio €[22]

The share of benefits is about equal for men (50,6%) on the one hand and women and children on the other hand (47,4% and 1,94%).

Figure 1.4 shows insurance prices for women at the age of 20, 30, 40, 50, 60 and 70, respectively. As only some insurances offer contracts for an entry age above 70, these prices are not shown here. One can see that not only prices rise with age, but that the price variance is increasing, too. While 75% of the offers are below monthly premium of 1€ per 5€ daily coverage at the age of 20, prices range from 5€ to 21.50€ at the age of 70.

These facts are even more obvious in Figure 5 that shows the corresponding box plots. The median price rises from 0.52€ for a 5€ daily coverage at the age of 20 to 9.05€ at the age of 70. The box area – and thus the range of the middle 50% of prices – is increasing with a rising age. Although there are no outliers besides one at the age of 20, the range of the prices is 7.6 times higher at the age of 70 than at the age of 20.

[22] PKV (2006), p. 49; PKV (2005), p. 85; PKV (2004), p. 87.

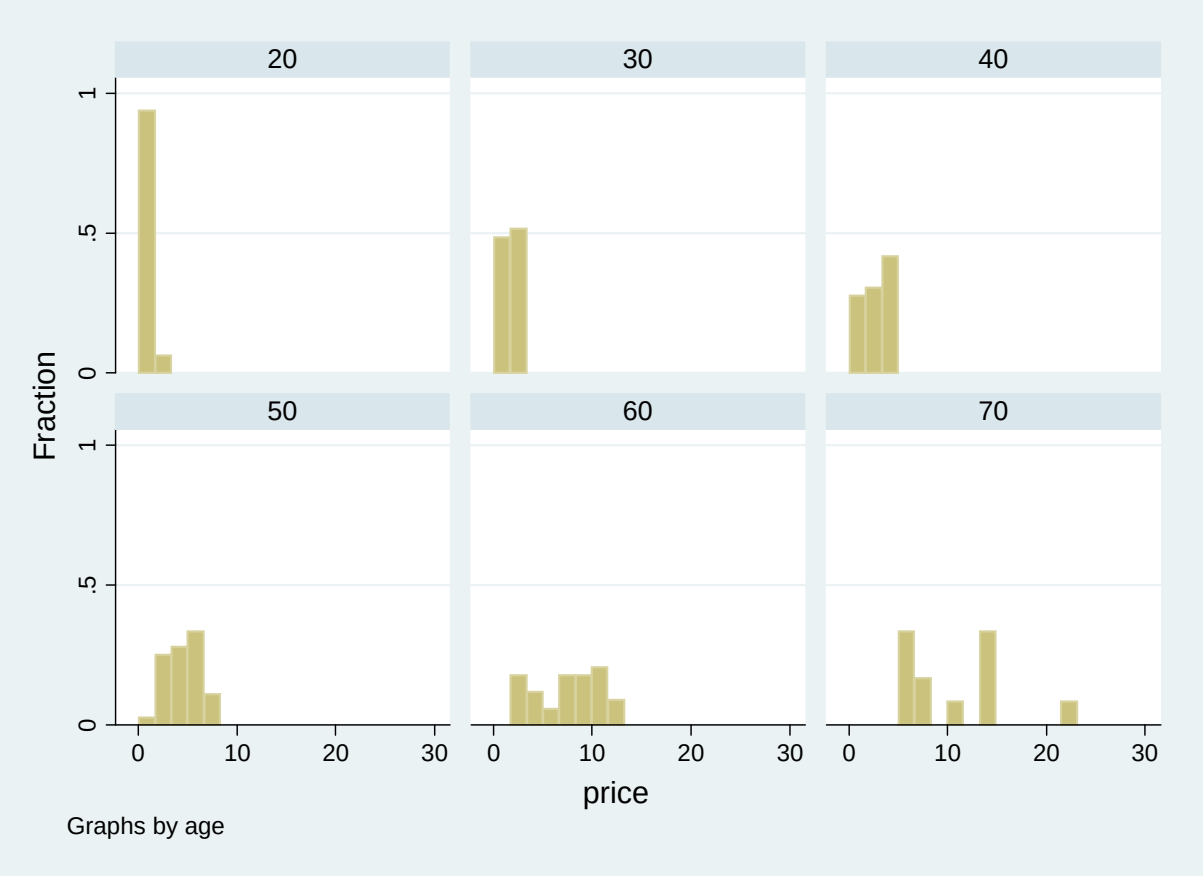

Figure 4.4: Histogram for Female Premiums by Age

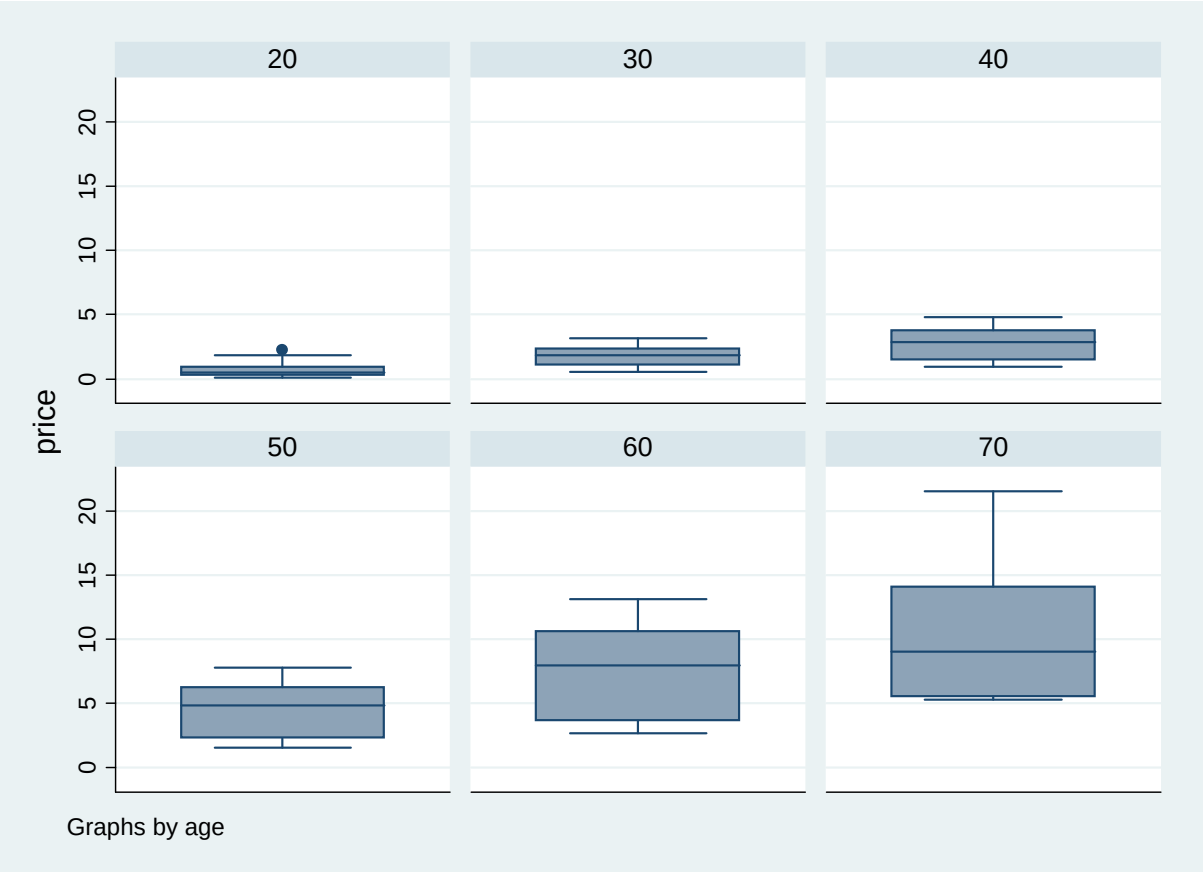

Figure 1.5: Box Plot for Female Premiums by Age

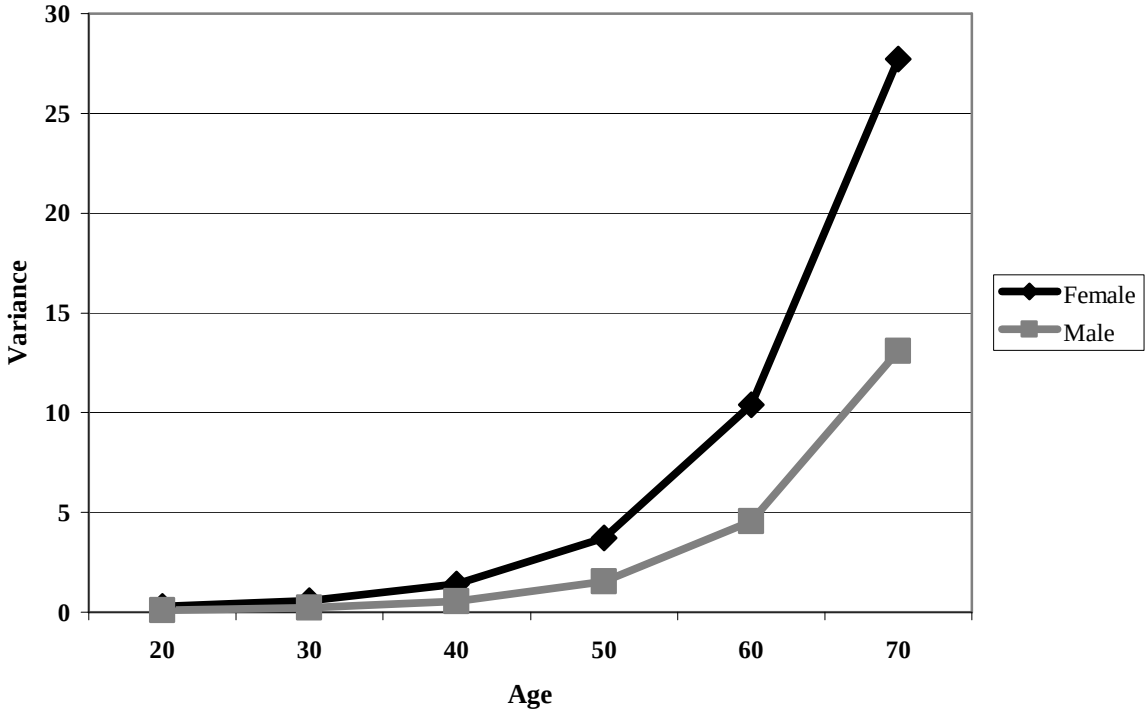

Figure 1.6: Price Variance for Female and Male Premiums

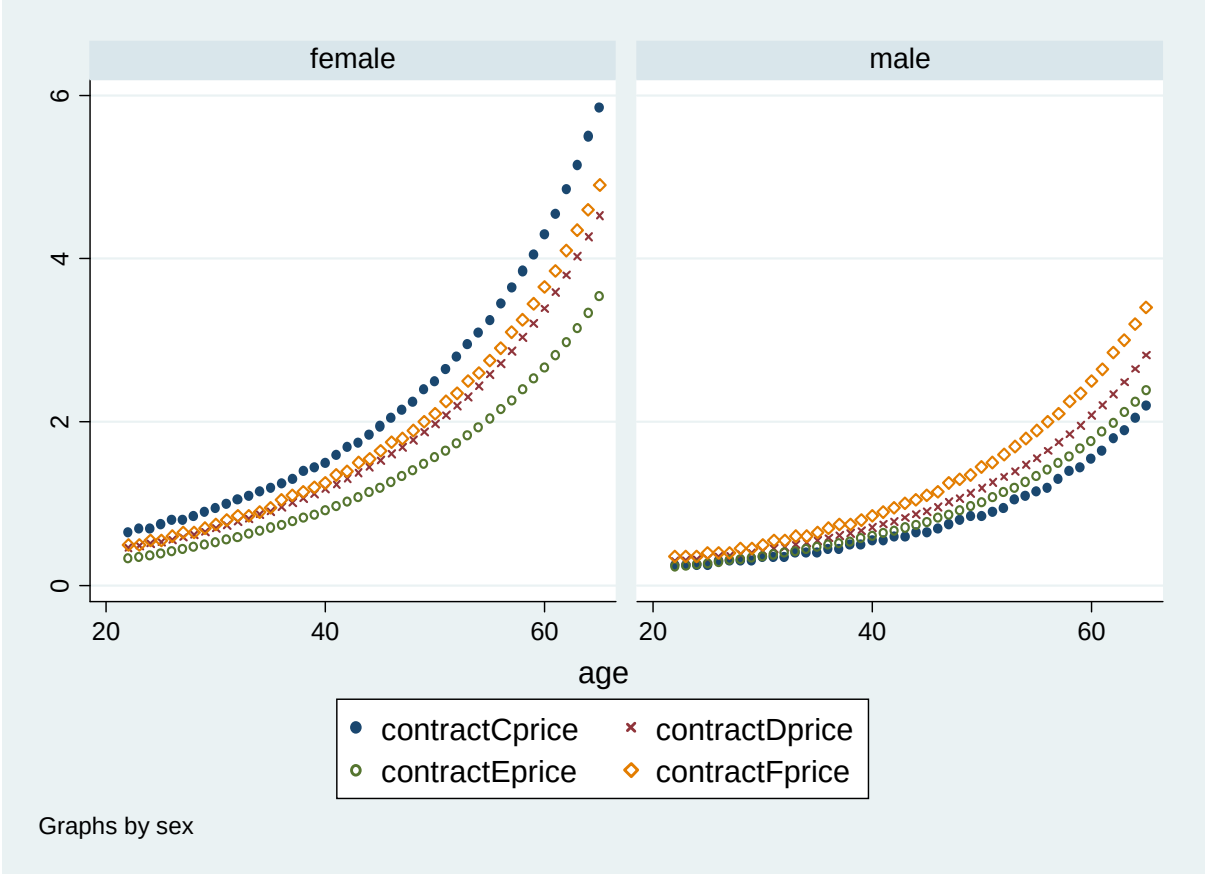

Figure 1.7: Price Comparison Firm C, Firm D, Firm E, Firm F

Results for men are mostly similar but price level and variance are less pronounced. Median prices rise from 0.45€ per 5€ coverage at the age of 20 to 6.02€ at the age of 70. The range of prices is 8.7 times higher at the age of 70 than at the age of 20.

Although the price variance is rising for both sexes, the effect is much stronger for females: as to be seen in Figure 1.6, the premium variance for women (men) increases from 0.30 (0.09) $€^2$ at the age of 20 to 27.73 (13.09) $€^2$ at the age of 70.

Price variance is still large if similar contracts are compared. Figure 1.7 shows four contracts that are identical with respect to the main contract characteristics like payments in varying care levels and the maximum age at entry. Not only are prices for identical products different, but the contract that is the most expensive for females is the cheapest contract for males.

1.4.3 Hedonic Regression

A hedonic regression is an OLS regression with the price of a product as the dependent variable and the product's characteristics as the independent variables. The characteristics' coefficients can be interpreted as marginal prices for one unit of the characteristic.

1.4.3.1 Premium Analysis

In the analysis of current insurance premiums the price of the insurance plan is the dependent variable. As the development of the premiums over the age at entry is clearly exponential, the price's logarithm is used as the dependent variable. Table 1.7 shows the variables in the regression and their descriptive statistics.

Variables	Mean	Std	Label
accidentwait	0.56	0.50	1 if standard waiting time does not apply in case of an accident
adultadjust	0.83	0.38	1 if premium adjusted at adult age (e.g. 16 or 21)
age	41.35	16.33	age at the moment of buying the insurance
age_gender	20.67	23.68	Interaction of age and gender: equals the age for men and zero for women
alwayswait	0.00	0.00	1 if there is a standard 3 year waiting time
carepremium	0.79	0.41	1 if premium payments in LTC
informalcare	0.16	0.37	1 if benefit adjustment for informal care taking
level1_euro	0.76	0.82	€ paid in care level 1
level2_euro	2.28	1.61	€ paid in care level 2
level3_dummy	0.30	0.46	1 if only care level 3 triggers payments
lprice	0.69	1.05	logarithm of the price for a premium with 5€ daily allowance in care level 3
max_age	67.21	8.30	maximal age at entry to contract the insurance
neverwait	0.32	0.47	1 if there is no waiting time between contracting and receiving benefits
older	38.90	20.49	age if age>=22, 0 else
sex	0.50	0.50	1 if male, 0 if female
young	0.14	0.35	1 if age at entry <22

Table 1.7: Variables and Values (3676 Observations)

	(1)	(2)
agespl0	0.121***	
	(0.0097)	
agespl21	0.067***	
	(0.0054)	
agespl30	0.0408***	
	(0.0019)	
agespl40	0.0472***	
	(0.00072)	
agespl50	0.0506***	
	(0.00056)	
agespl59	0.0361**	
	(0.014)	
agespl60	0.057***	
	(0.0025)	
agespl64	0.0202	
	(0.015)	
agespl65	0.0468***	
	(0.0093)	
agespl70	0.0448	
	(0.031)	
agespl75	0.0376	
	(0.037)	
agespl80	0.0799***	
	(0.0032)	
agespl90	0.106***	
	(0.0014)	
gender	0.227***	0.205***
	(0.043)	(0.045)
age_gender	0.00474***	0.00525***
	(0.00093)	(0.001)
degree1_euro	0.185***	0.185***
	(0.051)	(0.051)
degree2_euro	0.161***	0.162***
	(0.041)	(0.041)
degree3_dummy	-0.292**	-0.292**
	(0.13)	(0.13)
accidentwait	-0.0177	-0.0167
	(0.073)	(0.073)
neverwait	-0.0246	-0.0234
	(0.08)	(0.079)
max_age	0.00733***	0.00712***
	(0.0024)	(0.0022)
informalcare	0.0147	0.0166
	(0.083)	(0.082)
adultadjustment	-0.0992	-0.0996
	(0.078)	(0.077)
carepremium	-0.0334	-0.0316
	(0.046)	(0.046)
young		0.116
		(0.096)
older		0.0334***
		(0.0038)
older_sq		0.000153***
		(0.000036)
Constant	-3.888***	-1.91***
	(0.29)	(0.22)
Observations	3,676	3,676
R-squared	0.94	0.93

Table 1.8: Hedonic Regression Results
(with age splines in (1); dummy for young in (2);
standard errors in brackets, ***/**/* for significance at 1%/5%/10% level)

As insurance prices are recorded for each age at entry and both sexes, there are about 120 observations for each contract, depending on the maximum age at entry. Standard errors are thus clustered by contracts.

I present two regression models that differ in the treatment of the age structure in the data.

The first model uses a spline curve to identify a step-wise semi-logarithmic regression function. The knots are set as follows: As insurance prices are usually constant for entering children and young persons up to 21 years, the first knot is set at the age of 21. Then I use ten year steps up to the age of 59 where the contract with the smallest maximum age at entry has its upper bound. The knots at the ages of 60, 64, 65, 70 and 75 represent upper bounds for maximum ages at entry to the contracts as well. Up to the age of 100 years, I resume the ten year step structure.

The second model introduces a dummy for persons below the age of 22. The variable "older" is the age at entry above and equal to 22. To allow for a growing influence of the age with rising age at entry, an age-square variable "older_sq" is included in the regression.

An interaction effect between gender and age at entry is included in both models to account for a changing influence of the age in both sexes.

The regression results are reported in Table 1.8. The first column shows the results for the model with age splines. 94% of the price variation can be explained with this model.[23] The second model is shown in column (2). With this regression (model (2)), 93% of the price variation can be explained.

Age at entry is the most important factor for the premium. The influence of the age rises with the age, reflected by the rising positive significant coefficients of the age splines in model (1) and by positive significant coefficient of the squared age. As in model (2) most of the age information is covered by "older" and "older_sq", the young dummy is not significant.

Prices rise with age at entry and are 22.7% (20.5%) higher for women than for men in model (1) (model (2)). With rising age this effect is getting stronger as the coefficient of the interaction term between gender and age is positive and significant at 0.005. With rising age at entry the premium increases more for women than for men.

Payments in care levels other than care level three raise the insurance price about 18.5% for an extra Euro per 5€ payment in level three. Clearly, contracts which limit payments to care level three are cheaper than those with payments in all levels. On average, a contract which only pays in care level three is 29% cheaper than other contracts. Prices rise with maximum

[23] For comparison: in a hedonic regression for personal computers, the German federal statistical office reached between 72% and 80% (Linz and Eckert (2002), p. 862).

age at entry as well, reflecting the growing risk of becoming a nursing case with rising age. One explanation for this is that insurance companies that accept elder persons who face a much higher risk of becoming a nursing case might concentrate on this focus group. They might thus have less strict risk examinations to allow the elderly to enter. The consequence is a worse risk pool for the whole insurance collective and thus higher premiums.

The dummies for different waiting times are not significant. This reflects that these features of the insurance contract do not influence the premium calculation much.[24] As waiting times are standard features of health insurance contracts, this feature might just have been transferred to the SLTC insurance conditions. There are dummies for three further characteristics of the insurance contracts in the model: "informalcare" is equal to one if the insurance pays less for persons receiving care from informal personnel, e.g. members of the family. This dummy's influence is not significant. Neither is the dummy "adultadjustment" indicating whether premiums have to be adjusted at a certain age, e.g. the age of 21. The third dummy equals one if insurance premiums still have to be paid in case of LTC. Even if the coefficient of this variable is appropriately negative – if there is a premium in the case of LTC, premiums in the healthy case can be lowered – this coefficient is not significant either. Still, as these three variables are important characteristics of SLTC insurance contracts and potentially influencing a customer's choice between contracts, they are included in the regression.
According to model (2), a sample contract for a man born in 1964 entering the contract in 2007, including 20€ payment in care level one and 35€ in level two, without waiting time, with a maximum age at entry of 65, 100% premium payment for an informal care giver, premium adjustment at the age of 21 and no further premium payments in the case of LTC would cost 29,65€ per month for a daily allowance of 50€ in case of care level three. A woman of the same age at entry would pay 45,62€ for the same contract. The price difference reflects the difference in the probability of becoming a nursing case.
Even if most of the contract-specific characteristics in the model like a standard waiting time or benefits for informal care givers are insignificant, model(1) can explain 94% of the price variation in insurance contracts. Age and gender alone account for 74% of the price variation.

1.4.3.2 Firm Dummy Analysis

To analyse the market structure, I use dummies for the firms which offer SLTC insurance policies and include them in the regression. Many insurers offer only one contract. These firm

[24] A conversation with an actuary confirmed this view.

dummies, together with the dummies for contract specifics, cause problems of collinearity because the firms are already identified by the contract characteristics. This is why I exclude three non significant contract characteristics from the regression. As the reference for the dummies is the insurer with the largest number of customers in the health insurance sector, the coefficients show the price difference of the competitors to the market leader in percent. Results, again for model(1) and model(2), are shown in Table 1.9.[25]

	(1)	(2)
Firm1	-0.444***	-0.442***
	(0.056)	(0.057)
Firm2	-0.284***	-0.283***
	(0.031)	(0.031)
Firm3	-0.177***	-0.178***
	(0.045)	(0.046)
Firm4	0.105**	0.108**
	(0.042)	(0.043)
Firm5	-0.0476	-0.0462
	(0.09)	(0.09)
Firm6	-0.16	-0.156
	(0.095)	(0.096)
Firm7	0.0542*	0.053*
	(0.031)	(0.031)
Firm8	-0.00881	-0.0121
	(0.028)	(0.027)
Firm9	-0.207***	-0.207***
	(0.061)	(0.061)
Firm10	-0.0803	-0.0795
	(0.057)	(0.057)
Firm11	-0.364***	-0.365***
	(0.055)	(0.057)
Firm12	-0.0487**	-0.0479**
	(0.022)	(0.022)
Firm13	-0.228***	-0.229***
	(0.053)	(0.054)
Firm14	-0.477***	-0.475***
	(0.02)	(0.02)
Firm15	-0.113**	-0.113**
	(0.046)	(0.047)
Firm16	-0.345***	-0.346***
	(0.037)	(0.039)
Firm17	-0.229***	-0.229***
	(0.047)	(0.047)
Firm18	-0.374***	-0.372***
	(0.063)	(0.063)
Firm19	-0.0715	-0.0717
	(0.049)	(0.048)
Firm20	-0.372***	-0.373***
	(0.042)	(0.043)
Firm21	-0.354***	-0.356***
	(0.041)	(0.042)
Firm22	-0.153***	-0.153***
	(0.044)	(0.045)
Firm23	0.177***	0.179***
	(0.019)	(0.019)
Constant	-3.867***	-1.89***
	(0.28)	(0.18)
Observations	3,676	3,676
R-squared	0.95	0.95

Table 1.9: Firm Dummies
(other variables not shown here;
standard errors in brackets, ***/**/* for significance at 1%/5%/10% level)

[25] Only firm dummies shown here.

The graphic analysis of the dummy coefficients gives an even clearer view over the market: The box plot for the firm dummies in Figure 1.8 shows that most of the insurers offer cheaper contracts than the market leader does. The median is at -0.1685 which means that the median firm in the market offers a contract that is 15.5% cheaper than the market leader's offer. The cheapest firm even offers a 37.9% lower price for a contract. 50% of the firms lie within a range of -0.3495 to -0.0485 which is between 29.5% and 4.7% below the market leader's price.

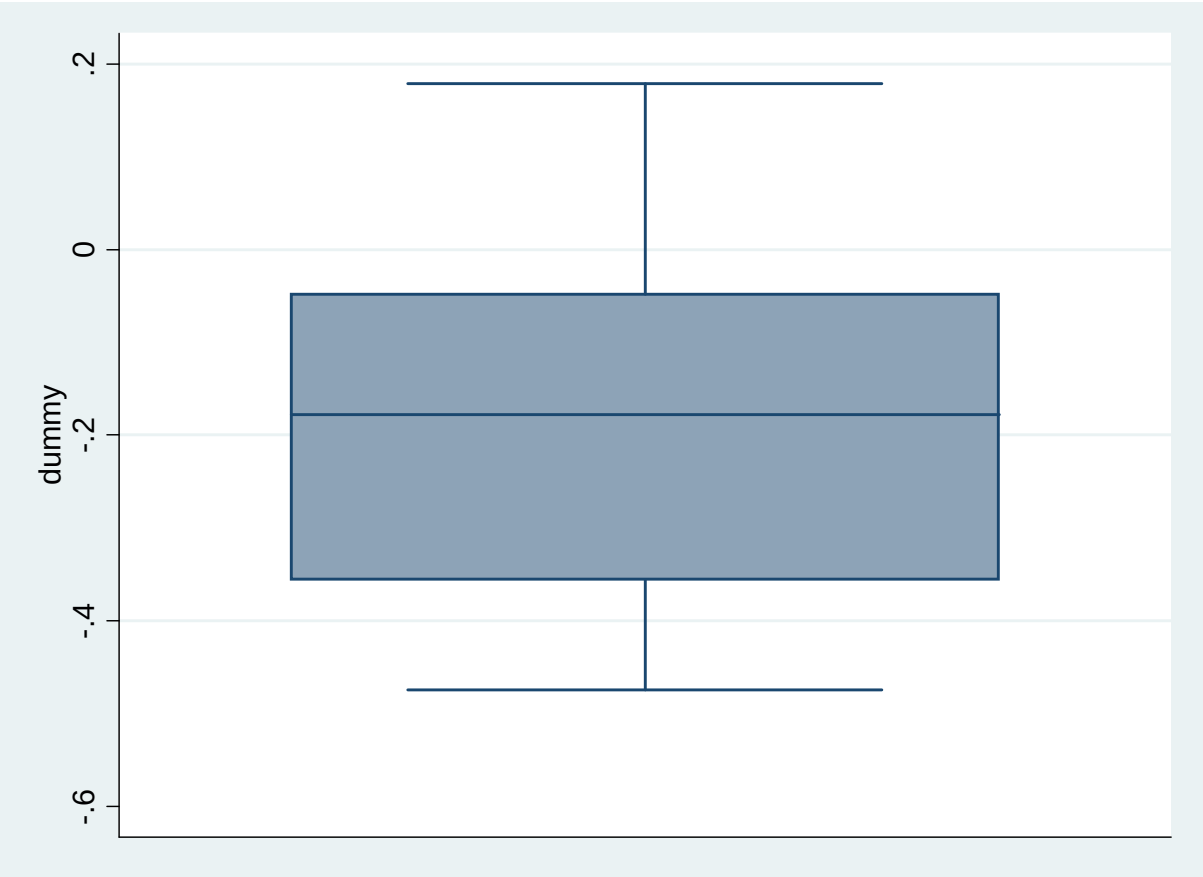

Figure 1.8: Box Plot for Firm Dummies

In Figure 1.9 and Figure 1.10, a comparison of the leading firms with the rest of the market is possible. To ease interpretation, the coefficients are already transformed into their percentage effects on the price. In Figure 1.9, the firms with the largest number of clients are compared to the market leader. As this ranking is based on performance in the general health insurance market and as not all health insurers publish SLTC insurance rates, the eighth firm is missing in the Figure.

The ninth company offers the cheapest contract compared to the market leader, which is also the cheapest contract in the overall comparison with -37.93% (see Figure 1.10). The tenth company is the most expensive one, with a 19.36% higher price than the market leader. Fig-

ure 1.10 shows all the dummies in the market, but arranged according to size. Black columns stand for top ten companies. The cheapest two firms and two out of the three most expensive firms belong to the top ten companies in the health insurance market.

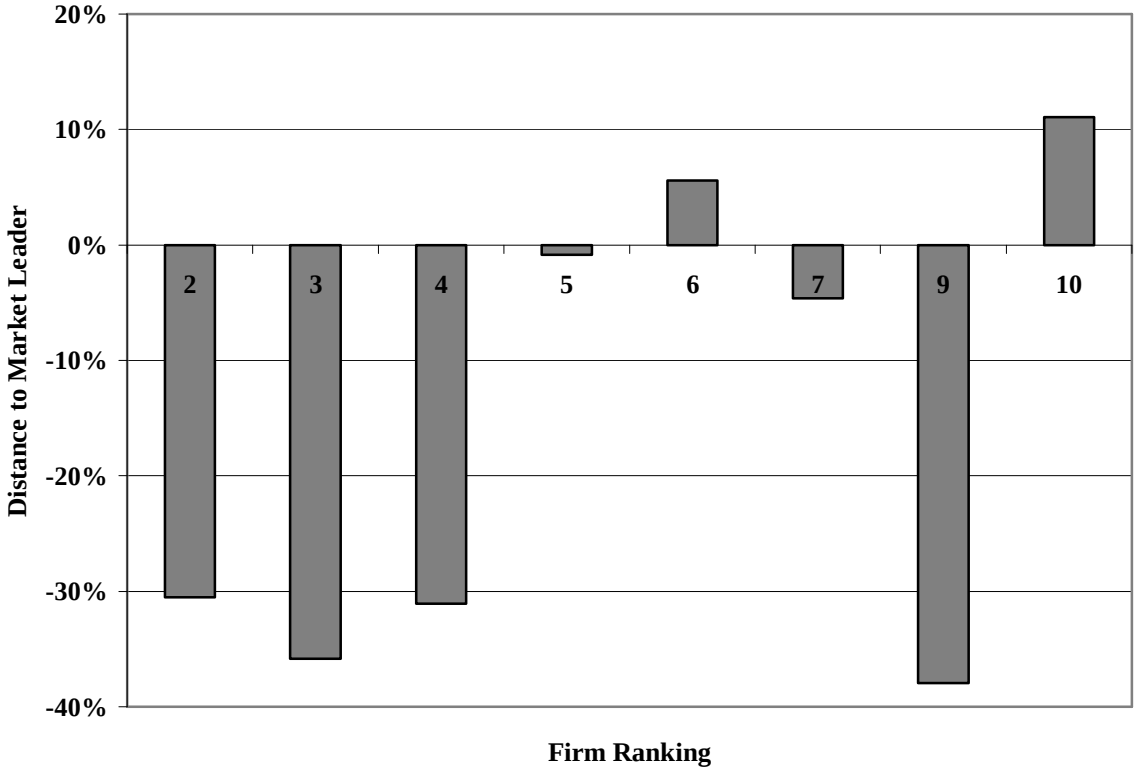

Figure 1.9: Top Ten compared to Market Leader

Insurer	Insured Persons	% of 8.36 Mio Insured Persons in Germany
Debeka	1983,006	23.7%
DKV	819,019	9.8%
Allianz Private Krankenversicherung	769,300	9.2%
SIGNAL	502,080	6.0%
DBV-Winterthur	434,251	5.2%
CENTRAL	405,279	4.8%
Bayerische Beamtenkrankenkasse	370,701	4.4%
Continentale	363,502	4.3%
HUK-Coburg	307,575	3.7%
Barmenia	299,325	3.6%
Total	6254,038	74.7%

Table 1.10: Insurer Ranking (2005)[26]

Another way to compare the top ten companies to the rest of the market is a split box plot as shown in Figure 1.11. In this Figure, the market leader is explicitly included with a dummy coefficient of zero on the top ten side. The median company of the top ten group offers at a higher rate than the median of the other group and the variance of the central 50% is higher than for the remaining firms. However, the means of both groups do not differ much with -0.173 in the top ten group and -0.177 in the rest group.

[26] Email from Stephan Caspary (PKV), 09-25-07.

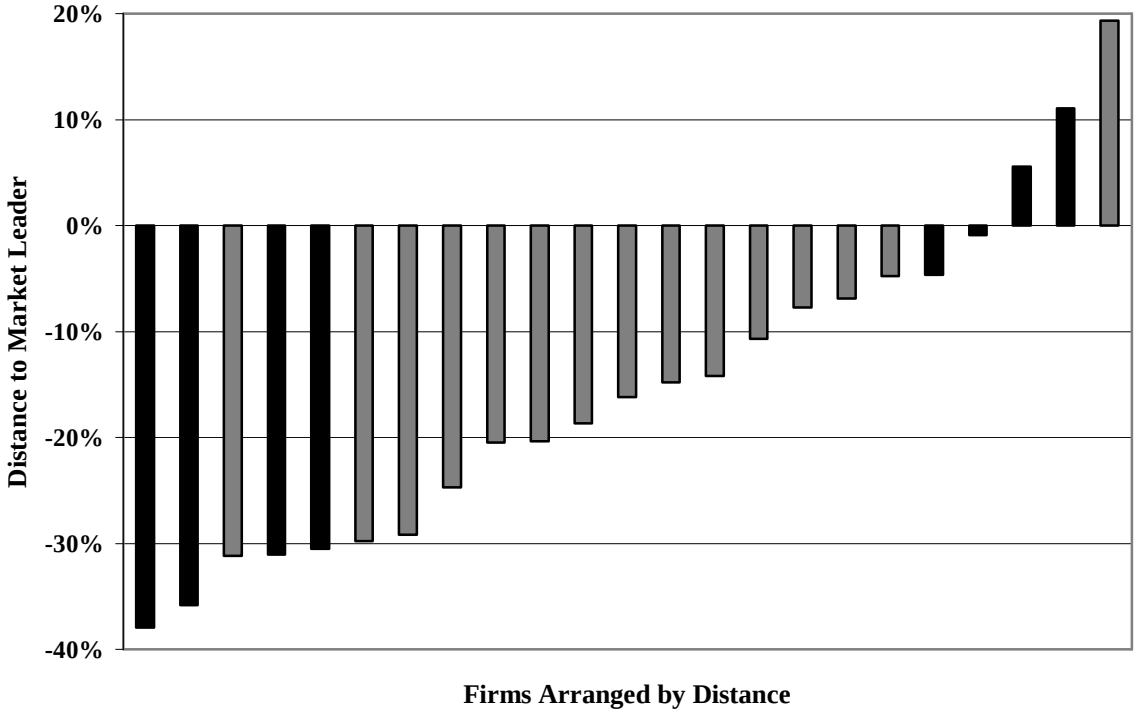

Figure 1.10: Top Ten compared to the Whole Market

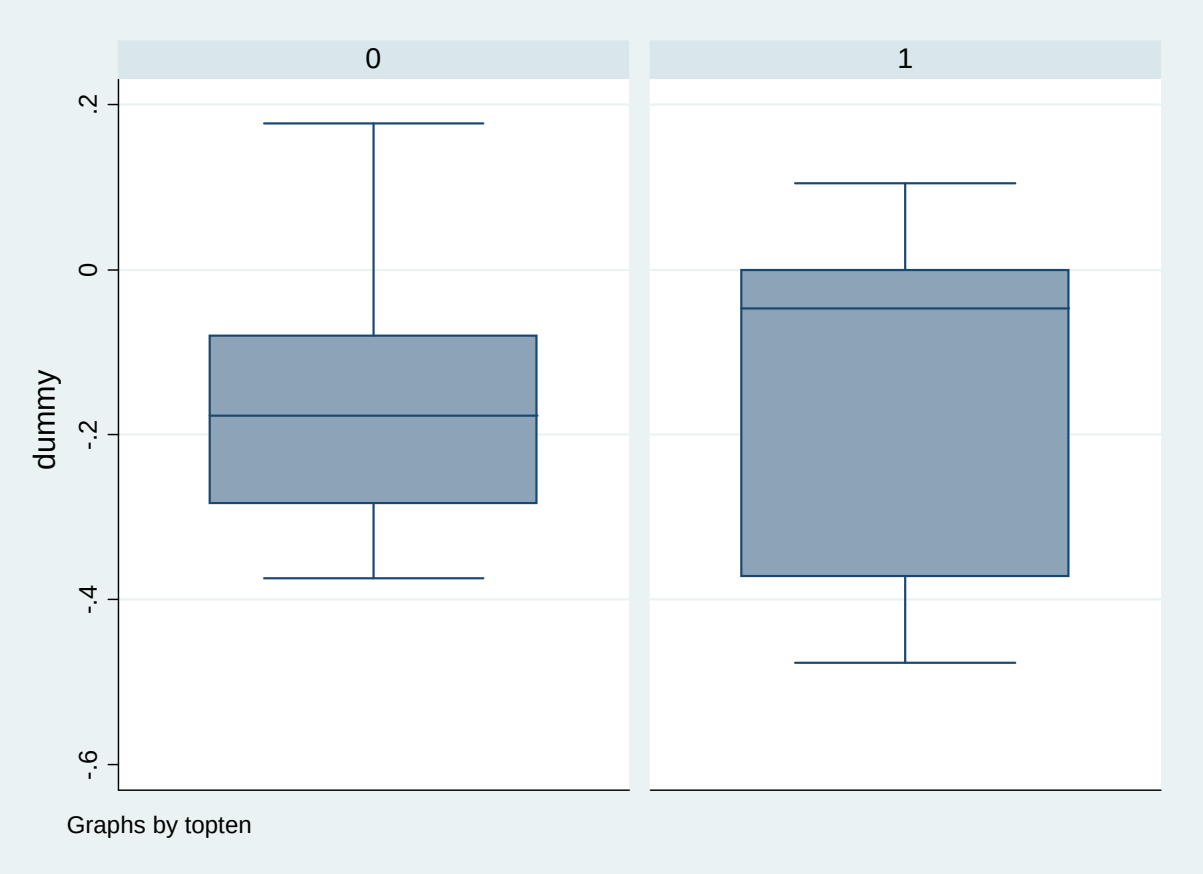

Figure 1.11: Box Plot for Firm Dummies by Top Ten Affiliation

1.4.4 Characterization of the Market

To characterize the market, several factors are important: the number of firms in the market, the price structure, the output and the access to information. Unfortunately, information is available only on the accumulated output of the market in Germany.[27] Still, the other factors can be explored based on the findings in chapter 1.4.2 and 1.4.3.

26 out of 53 firms in the market for private health insurance offer daily allowance insurance policies and every top ten company offers this insurance. Clearly there is no monopolistic structure in the market.

The price structure is very complex. Instead of one market price there are differences of up to 800% between the cheapest and the most expensive contract available. The firm dummy regression shows that even respecting the differences in contract characteristics, price differences are very large, e.g. 37.9% between the cheapest firm and the market leader's offers. Even quasi-identical contracts vary substantially in prices.

Only little information is available about SLTC contracts. Information structures such as web search engines – a common information tool for, e.g., automobile insurance – have not been implemented for this kind of insurance. It took several weeks to gather the data I used in this analysis as a lot of insurers are quite reluctant to publish their rates. Therefore a price comparison is barely possible for a customer in the market and search costs are high.

This shows that the private SLTC insurance market is far from perfect competition. In Scharfenberg (2008) I confirm this view arguing that most market imperfections result in elevated prices and that the premiums in the market for SLTC insurance policies are indeed too high.

One possible explanation is that this market is very small with only 800.000 sold contracts. Insurers seem to concentrate on other businesses and to neglect this market. The fact that there has not been much publicity for SLTC insurance supports this theory.

Another explanation is that private health insurers often sell SLTC insurance in a bundle with different supplementary insurances like supplementary dental insurance or even pension funds that come with SLTC insurance as an "old age provision bundle".[28] Customers who own a private health insurance policy with one company very unlikely insure their LTC risk with another company, especially as search costs are very high.

[27] Email from Stephan Caspary (PKV), 10-31-07. See 4.2 for aggregate numbers.
[28] See, e.g., Allianz (2007).

1.5 Policy Implications

This analysis shows that competition in the market for SLTC insurance is not perfect. Furthermore, only 1% of the German population owe an insurance policy and 50% of private health insurers, who could sell such policies, do not offer SLTC insurance contracts. What are possible policy implications of these findings? Should the government intervene and regulate the market?

Before these questions can be answered, one has to consider whether SLTC insurance is necessary at all. Arguments for the welfare effects of private supplementary insurance in the health sector can be derived from Hansen and Keiding (2002) on the one hand and Kifman (2002) on the other hand, and Danzon (2002) who compares and merges the results of the above-named. Hansen and Keiding compare three health insurance regimes, namely an unregulated, voluntary market regime, a compulsory and uniform universal regime with community rating and a compulsory regime that permits private supplementary insurance. Their analysis is based on a median voter model for the level of compulsory insurance and they alternatively use Hicksian compensation schemes and a utilitarian average utility concept to compare the welfare effects of the three systems. They show that in a voluntary insurance scheme low risks are worse off as they do not have any coverage as compared to a basic coverage at reasonable cost in the compulsory scheme. High risks profit from the voluntary market equilibrium as they can buy their individual adequate coverage which is higher than standard compulsory coverage would be. This is where the advantage of supplementary insurance lies: as high risks can cover the risk which remains after compulsory coverage on an individual basis this scheme increases welfare as compared to the compulsory scheme alone. Even if Hansen and Keiding also show that a pure voluntary insurance market can be welfare superior to the compulsory scheme, this case is not relevant for this analysis, taken that the study's intention is not to reform the whole LTC insurance system in Germany but rather to suggest ameliorations with respect to the SLTC market.

At first, Kifman's results seem to contradict Hansen and Keiding. He finds that if supplementary insurance covers specific services which are especially interesting for low risks, e.g. health clubs, they induce a higher willingness to pay among low risks than among high risks. Insurers could then try to attract a disproportionate share of low risks with supplementary insurance and – assuming the same insurer offers compulsory and supplementary insurance – reach a separate equilibrium with low and high risks buying different policies from different suppliers. However, the German market for LTC and SLTC insurance is widely organized in a way that compulsory and supplementary insurance cannot be offered by the same firm. Most

Germans have compulsory LTC insurance with a public health insurer whereas private SLTC insurance can only be offered by private health insurers. Therefore the problem of insurers attracting only low risks for both kinds of insurances is limited to those who are privately insured.

Combining the two studies one can conclude that the welfare effect of a supplementary insurance strongly depends on who buys the policies. If it is the high risks, supplementary insurance is more likely to improve welfare, especially if it is offered by separate insurers.

In Germany, risk examinations prevent the highest risks from entering the SLTC insurance. However, a lot of high risks can enter at increased premiums and only past and current health status can be measured. Furthermore, healthy customers can still have private information about their elevated LTC risk and thus profit from private SLTC insurance. Therefore SLTC insurance can increase welfare from a theoretical point of view.

Moreover, demographic changes which forced the government to reduce retirement provisions and to encourage people to buy supplementary pension funds on an individual basis will almost certainly call for parallel development in other branches of social security.

Assuming that SLTC insurance is important and indeed increases welfare, the question is whether the state should intervene in the market. Problems behind the imperfection of the market are numerous.[29] Besides a lack of competition, there is a major information problem on the demand side of the market. 28% of the German population believe that the compulsory LTC insurance provides full cost coverage.

For that reason the first step for the government would be to spread information about the limited coverage of compulsory LTC insurance. The discussion about a reform of the whole LTC system mainly focuses on how to finance existing benefits – the fact that benefits are not enough to cover the LTC cost is not mentioned in the media. Information about limited coverage should then be combined with the possibilities to insure the LTC risk on an individual supplementary basis. The official LTC handbook of the German Ministry of Health[30] refers to the limited coverage but proposes SLTC insurance only in the FAQ section. Reform proposals of the German Council of Economic Advisors (GCEA) claim individual elements in the LTC insurance market which follow a funding principle but do not mention the existing possibility of SLTC insurance policies.[31]

[29] See Scharfenberg (2008) for an investigation of the limited size of the market for private SLTC insurance.
[30] Bundesministerium für Gesundheit (2006).
[31] Sachverständigenrat (2004).

If more people are aware of the problem and possible solutions, demand for SLTC coverage is likely to increase. Also, information on SLTC insurance premiums should be available on a more transparent basis to allow potential customers to compare premiums and benefits. These effects will enhance competition without explicit governmental intervention in the market.
An important argument against buying SLTC insurance is the uncertainty in the LTC insurance system. Various pending reforms make people – and insurance companies – insecure with respect to how the system will change and how the current SLTC insurance system can be integrated into a new LTC insurance system. As mentioned above, reform ideas proposed by the GCEA never touch the existing SLTC insurance market.
The current LTC reform allows public health insurers to act as a broker for private SLTC insurance[32] which makes it easier for the customers to buy policies and thus potentially enhances demand for SLTC insurance. However, the official press statement does not inform about this fact[33]. Moreover, a complete makeover of the LTC insurance system – forcefully demanded by the GCEA – has not been implemented yet.

Combining these arguments it is obvious that the primary task for the government is to inform about the problem, about possible solutions – and about the role of private SLTC insurance in potential reform projects. This increase in the amount of information can raise demand and enhance competition in the market.

1.6 Conclusion

This paper is the first to analyze the German market for private SLTC insurance. This market is growing at a fast pace: benefit payments grow at rates between 20% and 25%. However, the market is still very small with 1% of the German population holding an SLTC insurance policy.
Private SLTC insurance markets only exist in France, the US, the UK and Germany. As premiums are very high, demand is very low in the UK.

An analysis of the available insurance contracts in the market shows that the variance in premiums rises with the age and is generally higher for women than for men.
The variation in insurance prices can be explained up to 94% by a semi-logarithmic hedonic regression. Besides age, squared age, gender and their interactions, contract-specific compo-

[32] Bundesministerium für Gesundheit (2007b).
[33] Bundesministerium für Gesundheit (2007c).

nents play an important role for the pricing. Premiums rise with age, especially for women who pay significantly more than men. A rising maximum age at entry into the insurance increases the price as well as a limitation of benefit payments to care level three. Additional characteristics such as payments for informal care or continued premium payments in the case of LTC do not have significant influence on the premium price.
After including firm dummies to the analysis, differences between the companies in the market can be analyzed. Most of the companies offer cheaper rates than the market leader. However, the top ten health insurers are not more expensive than the rest of the market. Besides that, high search costs for possible customers indicate that the market is far from perfect competition.

The main policy implications for the government lie on the demand side of the market. It is important to circle information about the fact that LTC insurance only covers part of the arising LTC costs and about possibilities to insure the remaining costs. Furthermore, with reforms of the LTC insurance system pending, the government should inform about the future role of SLTC insurance to reassure current and potential SLTC insurance customers. An increase in the amount of information about SLTC insurance can effect an increase in demand and thus enhance competition in the market as more customers make competition more attractive to the suppliers. The current LTC insurance reform which allows public health insurers to act as an agent for private SLTC insurance in front of their customers is a step into this direction.

References

Allianz AG (2007): Der Allianz Schutzbrief für alle ab 55. http://www.allianz.de/produkte/personen/schutzbrief_55plus/steckbrief/index.html (accessed in December 2007).

Brown, Jeffrey R., Norma B. Coe and Amy Finkelstein (2006): Medicaid Crowd-Out of Private Long-Term Care Insurance Demand: Evidence from the Health and Retirement Survey. NBER Working Paper 12536.

Bundesanstalt für Finanzleistungsaufsicht (2007): Gesetz über die Beaufsichtigung der Versicherungsunternehmen (Versicherungsaufsichtsgesetz - VAG). http://www.bafin.de/gesetze/vag.htm (accessed in December 2007).

Bundesministerium für Gesundheit (2006): BMG-P-G500, Pflegeversicherung. http://www.bmg.bund.de/cln_040/nn_603392/SharedDocs/Publikationen/Pflege/g-500,templateId=raw,property=publicationFile.pdf/g-500.pdf (accessed in December 2007).

Bundesministerium für Gesundheit (2007a): Zahlen und Fakten zur Pflegeversicherung. http://www.bmg.bund.de/cln_040/nn_773096/SharedDocs/Download/DE/Themenschwerpunkte/Pflegeversicherung/Informationen/ZahlenFakten,templateId=raw,property=publicationFile.pdf/ZahlenFakten.pdf (accessed in December 2007).

Bundesministerium für Gesundheit (2007b): Das bringt die Reform der Pflegeversicherung. http://www.bmg.bund.de/cln_040/nn_604244/DE/Themenschwerpunkte/Pflegeversicherung/Pflegereform-17-10,param=Links.html#doc1173806bodyText2 (accessed in December 2007).

Bundesministerium für Gesundheit (2007c): Pressemitteilung Pflege. http://www.bmg.bund.de/cln_040/nn_604244/DE/Presse/Pressemitteilungen/Presse-4-2007/pm-17-10-07.html (accessed in December 2007).

Colombo, Francesca and Nicole Tapay (2004): Private Health Insurance in OECD Countries: The Benefits and Costs for Individuals and Health Systems. *OECD Health Working Papers*, (15), OECD, Paris.

Danzon, Patricia (2002): Welfare effects of supplementary insurance: a comment. *Journal of Health Economics*, (21), 923-926.

Eliot, George (1876): Daniel Deronda, ed. by Graham Handley. Oxford, 1984.

Finanztest (2006): Pflege für alle. *Finanztest*, (4), 68-78.

Häcker, Jasmin and Bernd Raffelhüschen (2004): Denn sie wussten, was sie taten: Zur Reform der Sozialen Pflegeversicherung. *Vierteljahrshefte zur Wirtschaftsforschung*, (73), 158-174.

Hansen, Bodil O. and Hans Keiding (2002): Alternative health insurance schemes: a welfare comparison. *Journal of Health Economics*, (21), 739-756.

Kifmann, Matthias (2002): Community rating in health insurance and different benefit packages. *Journal of Health Economics*, (21), 719-737.

Linz, Stefan and Gudrun Eckert (2002): Zur Einführung hedonischer Methoden in die Preisstatistik. *Wirtschaft und Statistik* (2002), 857-863.

Milbrodt, Hartmut (2005): Aktuarielle Methoden der deutschen Privaten Krankenversicherung. DGVFM: *Schriftenreihe angewandte Versicherungsmathematik*, (34).

OECD (2005): Long-Term Care for Older People. Paris, 2005.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2004): Jahresgutachten: 2004/05 "Erfolge im Ausland - Herausforderungen im Inland". 2004.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2006): Jahresgutachten: 2006/07 "Widerstreitende Interessen - Ungenutzte Chancen". 2006.

Scharfenberg, Veronika (2008): "Why is Enrolment in the Market for Private Supplementary Long-Term Care Insurance in Germany so low?". Working Paper.

Schulz, Erika, Reiner Leidl and Hans-Helmut König (2004): The impact of ageing on hospital care and long-term care – the example of Germany. *Health Policy*, (67), 57-74.

Statistisches Bundesamt (2007): PFLEGESTATISTIK 2005: Pflege im Rahmen der Pflegeversicherung. 4. Bericht: Ländervergleich – Pflegeheime. https://www-ec.destatis.de/csp/shop/sfg/bpm.html.cms.cBroker.cls?cmspath=struktur,vollanzeige.csp&ID=1019863 (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2000): Die private Krankenversicherung 1999/2000. Zahlenbericht. Köln 2000 http://www.pkv.de/downloads/Zb00.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2001): Die private Krankenversicherung 2000/2001. Zahlenbericht. Köln 2001 http://www.pkv.de/downloads/Zb01.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2002): Die private Krankenversicherung 2001/2002. Zahlenbericht. Köln 2002 http://www.pkv.de/downloads/Zb02.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2003): Die private Krankenversicherung 2002/2003. Zahlenbericht. Köln 2003 http://www.pkv.de/downloads/Zb031.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2004): Die private Krankenversicherung 2003/2004. Zahlenbericht. Köln 2004 http://www.pkv.de/downloads/Zb04.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2005): Die private Krankenversicherung 2004/2005. Zahlenbericht. Köln 2005 http://www.pkv.de/downloads/Zahlenb_04_05.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2006): Die private Krankenversicherung 2005/2006. Zahlenbericht. Köln 2006.
http://www.pkv.de/downloads/Zahlenb_05_06neu.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2007): Die Reform der Pflegeversicherung. Arbeitsversion. February.
http://www.pkv.de/downloads/Bericht_Pflegeversicherung_22feb07_komplett.pdf (accessed in December 2007).

2 Why is Enrolment in the Market for Private Supplementary Long-Term Care Insurance in Germany so low?

2.1 Introduction

Since 1995, everybody in Germany is automatically insured in compulsory long-term care (LTC) insurance. To cover the costs, employees have to pay 0.85% of their income plus 0.25% if they are childless. The insurance has been designed to cover basic risk – not full cost coverage.[34] Social LTC insurance is mandatory for every German citizen holding public health insurance. Persons holding private health insurance are obliged to insure their LTC risk with a private LTC insurance policy.

According to social insurance law a nursing case is defined as a person who, caused by physical, intellectual or mental disease or disability, needs a substantial provision of nursing care and assistance with her activities of daily living (ADL's). Based on daily duration and type of assistance, e.g. nutrition, personal hygiene or mobility, three care levels and a hardship case are defined as shown in Table 2.1:

	Care Level 1	Care Level 2	Care Level 3	Hardship Case
Assistance	≥ 1.5 hrs / day	≥ 3 hrs / day	≥ 5 hrs / day	>> 5 hrs / day
Basic Care (as part of assistance time)	≥ 0.75 hrs / day	≥ 2 hrs / day	≥ 4 hrs / day	>> 4 hrs / day

Table 2.1: Care Level Definitions[35]

A person can only be assigned to a care level after she or her legal representative has applied for the LTC benefits. The official medical service of the German health insurances will then send an agent to the person's home to assess the amount of assistance the nursing case needs and to decide about her care level. Benefit payments depend on the care level and on the characteristics of the care giver – see Table 2.2 for details.

		Care Level			
Type of Care		**1**	**2**	**3**	**Hardship**
At Home (Choice between two options, combinations possible)	Professional Help	384	921	1,432	1,918
	Daily Allowance for informal helper	205	410	665	
Nursing Home	Professional Help	1,023	1,279	1,432	1,688

Table 2.2: LTC Payments in €[36]

Today these benefits cover about 53% of the costs people in nursing homes have to face. In 1999, it was 58%.[37] As benefits are not adapted to inflation the real coverage of risk has de-

[34] Bundesministerium für Gesundheit (2007b), p. 7.
[35] Bundesministerium für Gesundheit (2007b), p. 15.
[36] Bundesministerium für Gesundheit (2007a), p.4.
[37] Statistisches Bundesamt (2007), p. 15; Bundesministerium für Gesundheit (2007a), p.5; Statistisches Bundesamt (2001), p. 13.

creased since the introduction of the Social LTC insurance. Due to the demographic structure the number of elder people in the country and thus the number of people in LTC will increase rapidly within the next 40 years.[38] As a consequence the funds available per nursing case will be reduced. Supplementary LTC (SLTC) insurance policies offer additional coverage for the LTC risk on an individual basis.

Insurance companies offer two types of SLTC insurance policies: the first type – "Pflegekostenversicherung" / care cost insurance – covers some of the actual costs that arise in LTC and that are not already covered by the compulsory LTC insurance. The second type – "Pflegetagegeldversicherung" / daily allowance insurance – defines a fixed amount of money as daily allowance which is paid in the case of LTC. This insurance is more flexible and thus more popular as benefits can also be used to pay relatives who take care of the insured person.

Despite existing possibilities to cover the LTC risk on a private basis only 830,000 persons in Germany have supplementary LTC (SLTC) insurance.[39] Why is the market for SLTC insurance so small?
First I describe the German market for SLTC insurance using general statistics for the supply side and a Probit regression for insurance demand. I find that age, sex, income and family characteristics influence the decision whether to buy an SLTC insurance policy. Health variables can not be shown to have a significant influence on this decision.
Clearly, some factors on the demand side of the market limit the number of customers. They are mainly problems of information: survey results show that only 72% of the Germans are aware that the benefit payments from the compulsory LTC insurance are not enough to cover all the arising costs. 41% are not worried about their future concerning LTC costs. 26% do not consider supplementary insurance necessary.[40]

In the main part of the paper I give evidence of supply side market failures that limit the size of the market following a model by Brown and Finkelstein (2007) who analyze the supply side of the US LTC market. According to their study most supply side imperfections in a market result in elevated prices and / or "quantity rationing", i.e. contracts that do not cover the whole risk. In Germany customers can freely choose the extent of coverage in their SLTC

[38] Schulz et al. (2004), p. 66.
[39] PKV (2006), p.31.
[40] PKV (2007), p. 7.

contract. Quantity rationing is not possible. Therefore I concentrate on insurance prices. In order to find out whether they are indeed elevated I analyze the premiums in the market and I calculate the so-called load factors on existing SLTC insurance policies. The load is the difference between one and the ratio of the expected present discounted values (EPDV) of the benefits and the insurance premiums. If the EPDV of payments for the insurance equals the EPDV of benefits from the insurance this load will be zero and called a fair load. If the load factors are very high, i.e. approaching one, premiums are elevated according to the EPDV model.

In order to compute the expected values I estimate probabilities for the transition into LTC. This is the first time probabilities for the transition into different care levels are estimated on the basis of a large German data set. The Mikrozensus Panel 1996-1999 is the first German panel data set which contains enough observations to identify transition probabilities between care levels. With the continuous real health state as the latent underlying variable I estimate an Ordered Probit model – split by gender – for the one-year transition between five ordered health states: perfect health, care level one to three and death.[41] The estimates of the regression are used to predict probabilities for every possible transition between states, death being an absorbing state. The predicted probabilities are stored in transition matrices for each gender and age group. They contain the transition probabilities for a one year transition from five states to five states.
Under the assumption that the current health state depends only on the state in the previous year the Markov property is fulfilled. Transition probabilities from any period to any period can thus be calculated using Markov chains. The resulting probabilities are consistent with existing studies about the LTC risk, mortality rates and their differences for both sexes.
Expected discounted values are calculated by using actual benefit payments and premiums from insurance offers I collected from 25 German health insurances, and discount rates of 4% and 6%, respectively.

My study results in three findings: First of all, even if premiums differ substantially between men and women the expected values are independent of gender. Besides that, an increase in the age at entry always leads to an increase of the load factor. This means that elder entrants receive fewer benefits compared to their premium payments than younger customers.

[41] The hardship case is defined on an individual basis so it can not be modelled easily. In addition to that, far less than 1% of the frail are hardship cases (see, e.g., Statistisches Bundesamt (2007a)). Therefore the hardship case is excluded from the analysis.

As the main result I find that even when various forms of insurer costs are taken into consideration prices in the market are elevated according to the EPDV model. These high prices can have their origin in supply side failures in the market, especially asymmetric information, imperfect competition and high transaction costs.

2.2 Literature Review

In Germany, various publications describe mandatory LTC insurance, its problems and possible solutions. Schulz et al (2004) predict utilization rates of LTC and find a dramatic increase in the number of LTC cases. Häcker and Raffelhüschen (2004) point out that the current LTC system substantially lacks sustainability and that existing ideas to reform the system will not be able to solve the problem. However, no analysis of the German market for private SLTC insurance is known to the author.

The main contributions to the literature about the market for private LTC insurance in the United States have to be attributed to Amy Finkelstein and varying co-authors. In the U.S., LTC expenditures have to be borne individually. Only after the individuals have met stringent asset and income tests, Medicaid is covering LTC costs. In 2004, LTC costs amounted for 8.5 percent of all health care spending in the United States and over one-third of Medicaid expenses were devoted to LTC. Despite the possibility of insuring the LTC risk on a private basis, only 10 percent of the elderly have private LTC insurance and only 4 percent of LTC costs are borne by private insurance.[42]
Finkelstein's studies can be divided into three main topics: an analysis of the limited size of the private LTC insurance market, studies about asymmetric information in the market and the interaction of public (Medicaid) and private insurance in the market for LTC insurance.

The so-called "Medicaid Crowd-Out" of private LTC insurance demand has been analyzed in Brown and Finkelstein (2004b) and Brown, Coe and Finkelstein (2006). The crowd-out effect is based on the fact that Medicaid is a payer of last resort which covers LTC costs only after any private insurance and individual income and assets have been taken into account.
In order to test the hypothesis that Medicaid as a secondary payer hinders demand for private LTC insurance Brown and Finkelstein (2004b) develop a utility-based model of a 65-year old risk averse individual choosing an optimal inter-temporal consumption path under uncer-

[42] See Brown, Coe and Finkelstein (2006) for U.S. LTC expenditures.

tainty about LTC costs. Based on this model and based on common Medicaid rules they calculate the willingness to pay (WTP) for a private LTC insurance contract as the dollar amount of the utility difference between the optimal inter-temporal consumption paths with and without private LTC insurance, respectively. WTP values are found to be similar for both sexes and rising with assets. When the authors examine the effect of Medicaid on the WTP, they reach three results: first of all, Medicaid strongly constrains demand for private LTC insurance as even without supply side failures in the market[43] two thirds of the wealth distribution would not want to buy comprehensive LTC insurance. As fixing problems in the private LTC market would not increase demand in the presence of the current Medicaid system, changes in that system are necessary for the market to expand substantially. Second, the reason for this large crowd-out effect is that Medicaid imposes a very high "implicit tax" on the benefits of private LTC insurance. This means that for the median male, 60% of the benefits are redundant of benefits Medicaid would otherwise have paid. The third finding is that since Medicaid insurance is far from being comprehensive, most individuals face about 40 percent of uninsured expenditures in the case of LTC. Thus, welfare gains of being able to complement Medicaid coverage – which is not actually possible – would be substantial. Taken together, Brown and Finkelstein (2004b) show that public insurance can crowd out private insurance even if the public risk coverage is not comprehensive.

Brown, Coe and Finkelstein (2006) provide empirical evidence for the crowd-out effect using data from the 1996, 1998 and 2000 waves of the Health and Retirement Survey (HRS). They concentrate on the question how the Medicaid asset protection rules, i.e. how much individuals may keep before Medicaid coverage kicks in, affect private LTC insurance coverage among individuals between 55 and 69. They use rule variation across U.S. states and find significant evidence that more generous asset protection lowers insurance demand. According to their findings, a $10,000 increase in protected assets is related to a decrease in insurance coverage of 1.1 percentage points. However, even if this change is relatively high, most individuals would remain uninsured if protected assets were substantially decreased. This analysis complements the findings of Brown and Finkelstein (2004b) who showed Medicaid's implicit tax to account for a large portion of the lack of private insurance purchases. The empirical analysis shows that changes in asset disregards do not influence Medicaid's implicit tax much and thus do not have a large effect on private insurance coverage.

In Germany, the compulsory LTC insurance is paying independently from existing private SLTC insurance coverage and vice-versa so the above results cannot be applied or tested for

[43] See Brown and Finkelstein (2004a) and Brown and Finkelstein (2007) for supply side failures in the market.

the German market. However, even if a crowd-out effect like in the Medicaid system cannot appear in Germany there seems to be a psychological effect: as 28% of the Germans are convinced that the compulsory LTC insurance covers the LTC expenditures completely, they do not enter the market for private SLTC insurance.

Asymmetric information in the private LTC insurance market in the United States has most prominently been analyzed by Finkelstein and McGarry (2006), Finkelstein, McGarry and Sufi (2005) and Finkelstein and Poterba (2006). Each analysis widens the classical adverse selection szenario in another dimension.

Finkelstein and McGarry (2006) provide empirical evidence that there are two types of private information in the insurance market: risk type and risk preferences. They use data from the Asset and Health Dynamics of the Oldest Old (AHEAD) cohort of the HRS to define two groups of individuals who own private LTC insurance: those who believe that their risk to become a nursing case is elevated as compared to the risk the insurer expects – the classical adverse selection scenario – and those whose preference for insurance is above the average which the insurer observes. Finkelstein and McGarry show that the first group indeed has a higher LTC risk while the second group has below average use of nursing homes as this second group seems to comprise the more cautious individuals who are wealthier and/or who are more likely than average to also invest in health activities that the insurer cannot observe. In equilibrium, these two groups can cancel each other out so that on average the insured have a similar risk profile as the uninsured – even if, in fact, asymmetric information is present. As a consequence, the standard test of asymmetric information can lead to incorrect conclusions if insurance coverage and risk occurrence seem to be uncorrelated in the market. Therefore Finkelstein and McGarry propose an alternative test for asymmetric information which is robust to the existence of preference heterogeneity in insurance demand. Finkelstein and Poterba (2006) describe the test in detail. It is based on finding individual characteristics that are not used to price insurance policies, but that are correlated with insurance demand and / or with subsequent risk experience. An important quality of this new test is that in some cases it can identify adverse selection and not just generally asymmetric information. This is the case when external information points to unused characteristics that are correlated with risk type even when the insurance status does not differ across individuals, and when these characteristics take certain values for individuals who select more insurance. Even if Finkelstein and Poterba (2006) apply the test to the U.K. annuity market, results are applicable to the U.S. market for LTC insurance.

With the currently available data, this test cannot be applied for the German market for private SLTC insurance. First of all, the test requires information about the amount of insurance coverage and I only have data about whether someone is insured and not about how high coverage in the contract is. Furthermore, as the available data set only contains four subsequent periods, there is not enough time to observe the actual nursing home use of the insured individuals.

Finkelstein, McGarry and Sufi (2005) find another source of asymmetric information in a dynamic context: they amplify the definition of insurance benefits for risk averse individuals from a period-by-period "event" insurance to an insurance against becoming a high risk and being re-classified into a more expensive contract, i.e. an insurance against a "reclassification risk". The authors show that despite the theoretical possibility of constructing such insurance policies, the U.S. market for private LTC insurance does not insure this risk. The reason is a risk-based dynamic selection: after a while, those insurees who find out that they are of a lower risk than they thought before drop out of the contract. The effect is that the good risks leave the market and the bad risks stay. As a consequence, premiums have to be increased – and full insurance against reclassification risk is not possible. Based on the HRS, Finkelstein, McGarry and Sufi find that lapsing individuals are one-third less likely to subsequently be in a nursing home than those who stay with the contract – and that these results do not seem to be caused by moral hazard effects of keeping the insurance. However, they conclude that lapses are not always a consequence of updated beliefs about one's risk type. The data contains a high number of persons who cancel their policy one year after buying it – even if they are very unlikely to have found out more about their risk type this soon. Furthermore, as negative wealth or income shocks could account for a certain portion of the lapsed contracts the authors suggest exploring the empirical relevance of these other factors.

The German data set only contains four subsequent periods with information about insurance coverage. Unlike the HRS, there is no question about dropping out of the SLTC contract. Therefore I cannot distinguish between lapses and data problems. Thus, an analysis of lapsing behaviour is not possible.

The study that inspired my research is Brown and Finkelstein (2007) who answer the question why the market for LTC insurance is so small with an analysis of supply-side market failures in the market. They argue that supply side failures result in at least one of two empirical findings: elevated prices and "quantity rationing", i.e. contracts with limited coverage. In order to analyze prices, the authors calculate the "load" on contracts: the load is the differ-

ence between unity and the expected present discounted value (EPDV) of the insurance benefits over the EPDV of the insurance premiums. If the load equals zero the contract is "actuarially fair", meaning that, in expectation, people pay as much into the insurance as they are expected to be paid back. The lower benefits are as compared to premiums, the higher is the load factor. If the load is negative benefits exceed premiums. Brown and Finkelstein find that a contract typically bought by a 65-year old man has a load of 0.18 and that loads are substantially higher for men than for women. This is due to the fact that despite the much higher LTC risk for females premiums in the American LTC market are equal for both sexes. However, Brown and Finkelstein admit that neither high loads nor low benefits can fully explain the limited size of the market given that, e.g., insurance coverage is similar for both sexes even if load factors differ a lot.
I will refer to this article when I analyze the premiums in the German market for private SLTC insurance in chapter 5.

This study is the first to analyze transition probabilities between health, LTC and death based on the Mikrozensus Panel. A mathematical model to estimate transition probabilities between care levels in a German LTC portfolio has been developed by Czado and Gschlößl (2002) who use a semi parametric hazard model to analyze transitions between care levels. However, they do not have transitions between health and LTC and their data basis is much smaller.

2.3 The Market for private SLTC Insurance

In Germany, there are different ways to insure the risk of becoming a nursing case in a health insurance company. Private insurance companies offer two kinds of insurance. The first type is called "Pflegekostenversicherung" – care cost insurance. It covers certain costs that arise in LTC. This insurance is not very flexible and often not recommended.[44] As it pays only as the need arises and depending on a lot of circumstances, it is very hard to model. Furthermore, only 21% of the persons in private SLTC insurance actually own this kind of insurance.[45]
This is why this study analyzes the second kind of insurance, the "Pflegetagegeldversicherung" – daily allowance insurance (SLTC insurance in this paper) – which pays a certain amount of money in the case of LTC. This insurance is very flexible as the benefits do not depend on the project the money is being used for. Thus, frail persons can pay money to their relatives who take care of them. This daily allowance insurance is better to model as the pay-

[44] Finanztest (2006).
[45] PKV (2007), p. 31.

ment rules are alike for all insurance companies. Besides that this insurance is more popular in Germany (79%).

The SLTC insurance pays benefits irrespective of the payments of the compulsory LTC insurance and vice versa. There is no crowding out effect between the insurance plans as observed in the US by Brown et al. (2006). In order to avoid adverse selection, persons who buy an insurance policy have to be in good health. This is tested through a standard risk examination. In addition to that most of the contracts can only be closed if the applicant agrees to the condition that his family doctor may pass information about the applicant's health status to the contracting insurance company. If the applicant has severe health problems or diseases, insurance is either not possible or the premium is adapted to the elevated individual risk. As no data is available about these individually negotiated premiums, these cases cannot be taken into account in this study.

The market for private SLTC insurance is a small market compared to other private insurance markets.[46] In 2005 about 667.800 persons held daily allowance insurance, 8.8% more than in 2004.[47] The return from premiums was 170 Mio € when it was first recorded in 2005. Statistics before 2004 do not distinguish between daily allowance insurance and care cost insurance. For overall private SLTC insurance, the number of insured persons doubled from 380.000 in 1995 to 750.000 in 2003 and increased to 830.000 in 2005. Figure 1 illustrates this development:

Growth rates in benefit payments are even higher as more and more persons reach LTC age. Information is available back to 7.8 million € in 2002 and shows a 20-25% growth rate up to 14.9 million € in 2005. The raise in benefit payments of 22.7% from 11.9 Mio € in 2004 to 14.6 Mio € in 2005 was the highest increase in benefit payments in all private insurances in that period. Figure 2 shows the benefit development over the years 2002-2005.

Who are the 830.000 insured persons in Germany? In order to determine the characteristics of policy buyers, I run a Probit regression based on data from the Mikrozensus Cross Section 2003.[48] The binary dependent variable is whether someone owns an insurance policy. One percent in the sample has SLTC insurance – just like in the German population. The variables and their descriptive statistics are summarized in Table 2.3.

[46] See Scharfenberg (2008) for a detailed discussion of the limited SLTC market size.
[47] See PKV (2006), p. 31, 36, 49, 52, 86 for the following statistics.
[48] For details about the Mikrozensus see 2.5.2.1.

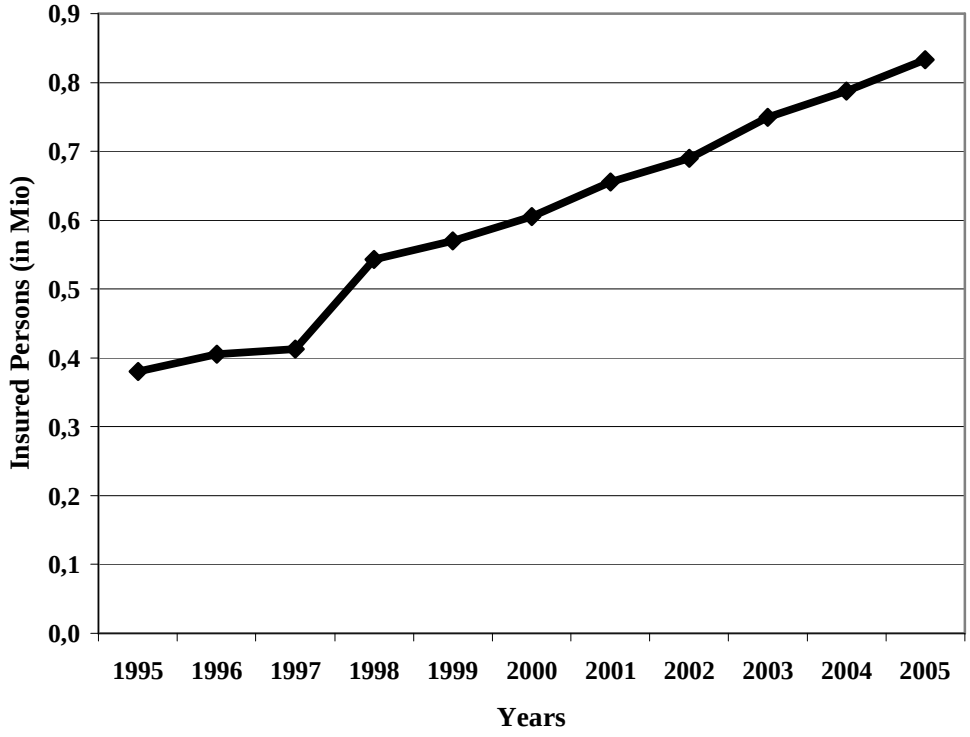

Figure 2.1: Insured Persons (in Mio)[49]

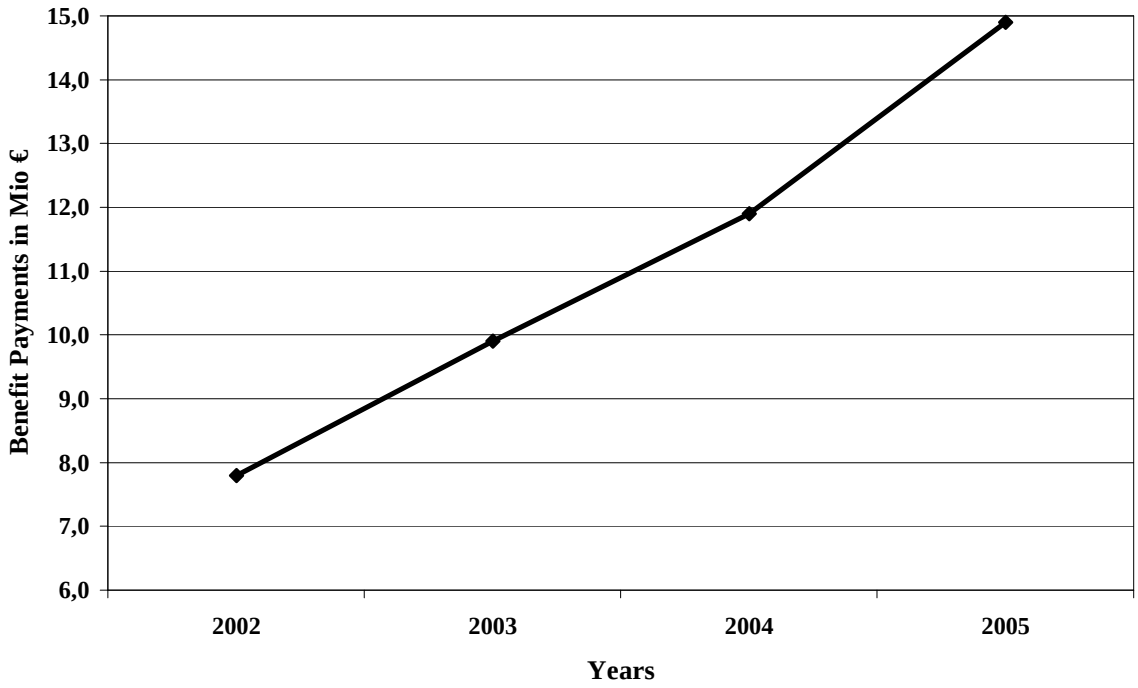

Figure 2.2: Benefit Payments in Mio €[50]

[49] PKV (2006), p. 86f.
[50] PKV (2006), p. 49; PKV (2005), p. 85; PKV (2004), p. 87.

	Observations	Mean	Std.Dev.	Min	Max
SLTC	493,035	0.011	0.105	0	1
Socio-economic Variables					
age	502,873	41.905	22.6074	0	95
sex	502,873	0.483	0.500	0	1
east	502,873	0.166	0.372	0	1
german	502,873	0.935	0.247	0	1
job	429,850	0.493	0.500	0	1
private	498,126	0.110	0.301	0	1
educ1	389,023	0.231	0.422	0	1
educ2	400,372	0.613	0.487	0	1
privatepension	130,844	0.085	0.279	0	1
income	480,633	968.811	1,186.404	0	18,000
Health Variables					
disabled	469,806	0.040	0.195	0	1
sick	469,780	0.044	0.206	0	1
accident	469,780	0.003	0.055	0	1
bmi	166,788	24.239	5.014	5.917	123.457
smoker	179,263	0.256	0.437	0	1
Family Variables					
kidnumber	497,953	0.911	1.149	0	11
kids	497,953	0.482	0.500	0	1
married	502,873	0.481	0.500	0	1
widowed	502,873	0.078	0.268	0	1
divorced	502,873	0.054	0.227	0	1
highcost	502,873	0.669	0.471	0	1

Table 2.3: Descriptive Statistics of the Probit Variables

The mean age in the sample is 41.9 years, 48.3 percent are male. 93.5 percent are German, 16.6 percent are from East Germany. 49.3 percent are currently employed and 11 percent of those who have health insurance own private health insurance. Education levels are defined by the highest degree of education a person has reached: educ1 equals one if a person has a general school diploma, educ2 identifies persons who have a higher degree like a university diploma. If both variables equal zero, the person has not graduated from any school. Privatepension defines persons who already provided for their old age on an individual basis. Income is the per capita household income defined as the total net household income adjusted by a standard equivalence scale.[51] Highcost is a variable that defines those Laender who offer LTC services at costs which lie above the German mean. Table 2.4 shows the variable highcost for the German Laender.

Probit regression results are shown in Table 2.5. The effect of age is positive and significant for nearly all the models as is sex. The older people get the more they have insurance. Men own more insurance policies than women but this effect weakens with the age. Persons from East Germany are significantly less insured than the rest of the country. If the variable highcost is included in the regression its coefficient is significant only if the dummy for East Ger-

[51] Income divided by the square root of household members.

many is excluded. This is due to the fact that no East German country has high costs and the highcost variable only mirrors the effect of East Germany.

Laender	Highcost
Schleswig-Holstein	1
Hamburg	1
Niedersachsen	0
Bremen	0
Nordrhein-Westfalen	1
Hessen	1
Rheinland-Pfalz	0
Baden-Württemberg	1
Bayern	1
Saarland	0
Berlin	1
Brandenburg	0
Mecklenburg-Vorpommern	0
Sachsen	0
Sachsen-Anhalt	0
Thüringen	0

Table 2.4: Laender with high LTC Cost[52]

The income coefficient is always positively significant and higher for males than for females. If income is split up into five dummies – the 10 percent poorest, 10 to 25 percent of the income scale, 25 to 50 percent, 50 to 75 percent, 75 to 90 percent and the 10% richest – the two lowest income classes are the less likely to have insurance and the richest are most likely to have bought a policy.

The health variables have significant influence on the insurance decision – the hypothesis that all health variables together equal zero can be rejected ($\chi_6^2 = 6.76$ in model (1)). However, smokers seem to be more likely to buy SLTC insurance as the coefficient is positive and significant in model (3).

Family variables like whether people are married or whether they have children do not influence the insurance decision. Only widowed and divorced persons can be shown to have less insurance. The fact that widowed men have more insurance than widowed women reflects the problem that usually women take care of their husbands who die earlier: if the woman dies first, the husband has to take care of himself. The influence of the education level does not differ between sexes. It is negative for low education levels and positive for higher levels.

Persons who already have private annuity insurance and those who have private health insurance are more likely to have SLTC insurance as well.

[52] Based on Statistisches Bundesamt (2007), p. 15.

	(1)	(2)	(3)	(4)	(5)
Socio-economic variables					
age	0.00459*	0.00517**	0.00384	0.00532**	0.00517
	(0.0024)	(0.0023)	(0.0024)	(0.0023)	(0.0023)
sex	0.383***	0.269*	0.341**	0.236	0.269
	(0.12)	(0.14)	(0.14)	(0.14)	(0.14)
agesex	-0.00848***	-0.00865***	-0.00607**	-0.00847***	-0.00865
	(0.0030)	(0.0026)	(0.0027)	(0.0026)	(0.0026)
east	-0.306***	-0.302***	-0.253***		-0.295
	(0.057)	(0.057)	(0.056)		(0.065)
german	0.189**	0.195**	0.149*	0.181**	0.196
	(0.087)	(0.087)	(0.088)	(0.088)	(0.087)
job	0.146***	0.151***	0.175***	0.149***	0.151
	(0.050)	(0.050)	(0.046)	(0.051)	(0.050)
private	0.335***	0.335***	0.287***	0.345***	0.335
	(0.044)	(0.043)	(0.043)	(0.044)	(0.043)
educ1	-0.0929	-0.179**	-0.120	-0.173**	-0.179
	(0.061)	(0.085)	(0.083)	(0.085)	(0.085)
educ2	0.0648	0.0144	0.0425	0.00829	0.0143
	(0.042)	(0.067)	(0.064)	(0.067)	(0.067)
sexeduc1		0.179	0.0590	0.191	0.179
		(0.12)	(0.11)	(0.12)	(0.12)
sexeduc2		0.0836	0.0457	0.0944	0.0837
		(0.084)	(0.078)	(0.084)	(0.084)
privatepension	0.111**	0.112**	0.111**	0.101**	0.112
	(0.049)	(0.049)	(0.048)	(0.049)	(0.049)
income	0.000103***	0.0000942***		0.0000936***	0.0000941
	(0.000029)	(0.000030)		(0.000030)	(0.000030)
income_sq	-6.79e-09***	-7.00e-09***		-7.43e-09***	-7.00e-09
	(1.85e-09)	(1.86e-09)		(1.88e-09)	(1.86e-09)
incomesex	0.0000327	0.0000462*		0.0000556**	0.0000463
	(0.000023)	(0.000026)		(0.000027)	(0.000026)
hhpoor			-0.224		
			(0.14)		
hhlow			-0.304**		
			(0.14)		
hhmiddle			-0.0224		
			(0.12)		
hhhigh			0.171		
			(0.11)		
hhrich			0.260**		
			(0.12)		
p-value (F-Test)	0.0000	0.0000	0.0000	0.0000	0.0000
Health Variables					
disabled	0.0322	0.0355	0.0338	0.0386	0.0356
	(0.055)	(0.055)	(0.056)	(0.055)	(0.055)
sick	0.0148	0.0143	0.0365	0.0131	0.0142
	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)
accident	-0.182	-0.182	-0.145	-0.180	-0.182
	(0.21)	(0.21)	(0.21)	(0.20)	(0.21)
bmi	0.0463	0.0477	0.0470	0.0469	0.0478
	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)
bmi_sq	-0.000939	-0.000962	-0.000904	-0.000963	-0.000964
	(0.00064)	(0.00064)	(0.00064)	(0.00065)	(0.00064)
smoker	0.0625*	0.0645*	0.0862**	0.0651*	0.0645*
	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
p-value (F-Test)	0.3434	0.3165	0.1986	0.2756	0.3165

…

Family Variables					
kidnumber	-0.0483				
	(0.033)				
kids	0.0741	-0.00720	0.00185	-0.00368	-0.00720
	(0.068)	(0.040)	(0.040)	(0.040)	(0.040)
married	-0.0474	-0.0727	-0.0555	-0.0718	-0.0725
	(0.078)	(0.053)	(0.054)	(0.054)	(0.054)
widowed	-0.311**	-0.325**	-0.222*	-0.326**	-0.325**
	(0.14)	(0.13)	(0.13)	(0.13)	(0.13)
divorced	-0.199	-0.241***	-0.188	-0.245***	-0.241***
	(0.12)	(0.084)	(0.085)	(0.084)	(0.084)
marriedsex	-0.0339				
	(0.097)				
widowedsex	0.367	0.385*	0.265*	0.374*	0.385*
	(0.23)	(0.21)	(0.21)	(0.21)	(0.21)
divorcedsex	-0.0751				
	(0.16)				
p-value (F-Test)	0.0469	0.0223	0.2264	0.0210	0.0225
highcost				0.114***	0.00912
				(0.036)	(0.042)
Constant	-3.286***	-3.267***	-3.207***	-3.355***	-3.276***
	(0.46)	(0.47)	(0.48)	(0.47)	(0.47)
Observations	34,181	34,181	34,048	34,181	34,181

Table 2.5: Probit Regression Results
(standard errors in brackets, ***/**/* for significance at 1%/5%/10% level)

2.4 Demand Side Limitations

Many theories explain the limited size of the US LTC market. Sloan and Norton (1997) conduct a literature survey on these theories finding that limitations can be caused by factors on the demand side or factors on the supply side of the market.

On the demand side, there are mainly two explanations for a limited market size: limited consumer rationality and the availability of imperfect but cheaper substitutes.

According to Kunreuther (1978), consumer rationality can be limited by the fact that people have problems understanding low-probability high-loss events. They tend to ignore these events, especially if these events have not occurred recently. The event of becoming a nursing case might fall into this category, especially for very young persons who have not been confronted with close relatives in LTC yet. Besides that, misconceptions about the compulsory LTC insurance system can substantially limit demand in the market: in Germany, 28 percent do not see a need for insurance thinking they are fully insured by their compulsory LTC insurance.[53]

Imperfect but cheaper substitutes can be financial help by children, informal unpaid care provided by family members (Pauly, 1990) or the compulsory LTC insurance. The latter currently covers 53 percent of LTC expenditures. It is cheap in a sense that premiums are compulsory and directly deducted from the pay check. No further individual payments are necessary.

[53] PKV (2007), p. 8.

The Probit regression in chapter 3 showed that persons with children are not more or less likely to buy SLTC insurance than childless persons – nor are married persons. This shows that Pauly's hypothesis that persons with children are less likely to buy LTC insurance as they want their family to take care of them is not valid here. People seem not to trust on their family as substitutes to cover their LTC risk.

Concluding, the fact that 28 percent of the Germans do not see a lack in LTC risk coverage is the most important argument for demand to be so small: as people see compulsory LTC insurance as a substitute for private insurance and as they think compulsory LTC insurance covers the total LTC costs, they do not buy SLTC insurance.

2.5 Supply Side Limitations: a Model to analyse Insurance Premiums

On the supply side, reasons for a limitation of the market size are less obvious. Supply side imperfections could be asymmetric information, imperfect competition, high transaction costs or the uninsured aggregate risks of rising LTC costs in the future. In order to determine failures in the market Brown and Finkelstein (2007) propose to analyze the insurance premiums in the market as most imperfections result in elevated prices and / or "quantity rationing", i.e. contracts that do not cover the whole risk.

In this section I apply the model by Brown and Finkelstein (2007) to find elevated prices in the market for SLTC insurance. I concentrate on prices because the insured decide themselves how much of their risk they want to have covered. Thus coverage is not constrained.

2.5.1 The model

Brown and Finkelstein (2007) analyze the supply side of the US LTC market calculating the "real price" of a contract, the load:

The load is the difference between unity and the expected present discounted value (EPDV) of the insurance payments over the EPDV of the insurance premium. If the load equals zero the contract is called "actuarially fair", meaning that, in expectation, people pay as much into the insurance as they are expected to be paid back. The lower benefits are as compared to premiums, the higher is the load factor. If the load is negative benefits exceed premiums.

$$\text{Load}_j = 1 - \frac{\text{EPDV(Benefits)}}{\text{EPDV(Premiums)}} = 1 - \frac{\sum_{t=j}^{100}\sum_{s=1}^{5}\frac{Q_{t,s}\cdot B_s}{(1+i)^{t-j}}}{\sum_{t=j}^{100}\sum_{s=1}^{5}\frac{Q_{t,s}\cdot P_{s,j}}{(1+i)^{t-j}}} \qquad (1)$$

The index t is denoted as the person's age. Purchase happens at age t=j where t goes from j to 100 in one year steps. Maximum life length is assumed to be 100 years here.

B_s is the benefit a person receives depending on the person's health state; P_s is the insurance price the person has to pay in her state. This price also depends on the age at purchase.

According to the model a person can be in five health states, deteriorating from 1 to 5: s=1 is in good health. A person in state 1 pays insurance premiums and does not receive benefits. As only individuals in good shape can contract the insurance, s=1 in t=j. s=2, s=3 and s=4 are the three care levels according to German law. I assume that people do not pay insurance premiums when in care but receive benefits.[54] s=5 is death. Clearly in this state payments and benefits are zero.

i is a realistic second-order discount rate. I use values of 4% and 6% here.[55]

Equation (1) shows the basic model I use to analyze the SLTC insurance market as a whole. Contract characteristics are not taken into consideration. Especially, I do not consider waiting times between buying the policy and receiving benefits. In my calculations for single insurers I will account for contract characteristics.

2.5.2 Probabilities of the Transition into LTC

The most important input factor is $Q_{t,s}$. This is the conditional probability of being in state s in time t given that the individual was in state 1 at the time of purchase. In order to calculate Q I use data from the Mikrozensus Panel 1996-1999.

2.5.2.1 The Mikrozensus Data

The Mikrozensus is a representative sample of the German population. Participation is enforced by law which leads to a very high participation rate. Once a year the interviewers conduct face-to-face interviews to ask the participants about their social and economic situation, their education and the job market.

Until 2006 the Mikrozensus data has only been published as cross-sectional data. This is the first time the panel structure of the data can be used. Every year one quarter of the panel is being replaced which leads to a four-year rotating panel with 800.000 participants in total. The Mikrozensus Panel 1996-1999, published in autumn 2006, can be used to estimate probabilities for the transition between care levels.

[54] For some insurers, premiums have to be paid in all living states. This fact cannot be implemented in the general analysis but will be in the single insurer analysis.

[55] These are values I collected speaking with an actuary: 4% is the lower bound, 6% the upper bound for calculations.

2.5.2.2 Calculation

As detailed information about the care level is only available for the first three years of the panel, no information from 1999 could be used. The panel starts in 1996 and no dead individuals are included in the 1996 data. To account for the fact that death is an absorbing state observations with death as the state in the previous year are excluded from the regression. As the number of observations for the health state is not sufficient for persons below the age of 20, I exclude them from the analysis. In the data, age is recorded only for persons born after 1900. Therefore the maximum age in the estimation is 98 years based on the transition 1997-1998.

In total the data set includes 89,757 observations for the state, out of which 87,912 are in state one, 554 in state two, 528 in state three, 194 in state four and 569 die. In order to calculate the transition probabilities the transition 1996-97 and the transition 1997-98 can be used.

As individual risk adaptations for applicants who do not pass the health exam can not be considered in this analysis, only healthy persons can enter the insurance in this model. Premiums then only depend on age and sex – no other information is being used to calculate the premium. Therefore my first step is to estimate the probability of a transition between the five states using an Ordered Probit regression with the state as the dependent variable and dummies for the lagged states and age as the independent variables. An Ordered Probit regression is suitable as the dependent discrete variable is ordered with categories which decrease in their value (state one is the healthiest, state five is the less healthy). The latent variable behind the process is the individual's "true health status". If this status falls below a certain threshold the next worse state is reached. In order to account for a possible nonlinear influence of the age I use a spline curve for the age with one age knot every five years. Probabilities are estimated separately for men and women as the hypothesis that a joint estimation leads to the same results as the separate estimation can be rejected in a Likelihood Ratio Test ($\chi^2_{18} = 42,67$). See Table 2.6 for results from the Ordered Probit Model.

As observations in the data do not suffice to predict probabilities for each age, I predict probabilities based on an artificial data set. This data set includes one observation for each age, gender and health state so that one transition matrix for each gender and age can be predicted. These transition matrices contain the transition probabilities for a one year transition from four living states to five states (including death). The probabilities in each line, i.e. from one state to the five other states, add up to one. As death is the absorbing state, the probability of

the transition from death into one of the four living states is manually set to zero and the probability of staying dead is set to 1.

	Male	Female
agespl20	0.0205	-0.0872
	(0.052)	(0.061)
agespl25	0.00841	0.0259
	(0.042)	(0.059)
agespl30	-0.0315	0.0176
	(0.039)	(0.049)
agespl35	-0.0124	-0.0137
	(0.041)	(0.044)
agespl40	0.0197	0.0381
	(0.041)	(0.040)
agespl45	0.0831**	0.0593*
	(0.035)	(0.034)
agespl50	-0.0348	-0.0269
	(0.029)	(0.030)
agespl55	0.0611**	0.0301
	(0.025)	(0.028)
agespl60	0.0412	0.0395
	(0.021)	(0.024)
agespl65	0.0357	0.0303
	(0.020)	(0.022)
agespl70	0.0597***	0.0132
	(0.019)	(0.020)
agespl75	0.00263	0.0248
	(0.024)	(0.022)
agespl80	0.0634**	0.0511**
	(0.027)	(0.023)
agespl85	0.0102	0.0343
	(0.035)	(0.025)
agespl90	0.0389	0.0520
	(0.084)	(0.043)
agespl95	0.537	0.00550
	(0.35)	(0.12)
stateone_lag	-2.479***	-2.776***
	(0.13)	(0.11)
statetwo_lag	-0.582***	-0.742***
	(0.15)	(0.12)
statethree_lag	-0.290**	-0.357***
	(0.15)	(0.12)
/cut1	0.393	-0.197
	(0.25)	(0.25)
/cut2	0.561**	0.180
	(0.25)	(0.25)
/cut3	0.763***	0.642***
	(0.25)	(0.25)
/cut4	0.847***	0.856***
	(0.25)	(0.25)
Observations	43012	46745
Pseudo R^2	0.2893	0.3783

Table 2.6: Results of the Ordered Probit Regression
(standard errors in brackets, ***/**/* for significance at 1%/5%/10% level)

Under the assumption that the current state only depends on the state in the year before the Markov property is fulfilled. Transition probabilities from any period to any period can thus be calculated using Markov chains. Based on 79 possible ages (j) at entry to the insurance,[56] transition probabilities have to be calculated for 79 starting points and (79-j) end points. As the individual is assumed to be healthy at entrance, i.e. to be in s=1 in t=j, only the transition

[56] 20 being the lower bound, 98 being the upper bound to age.

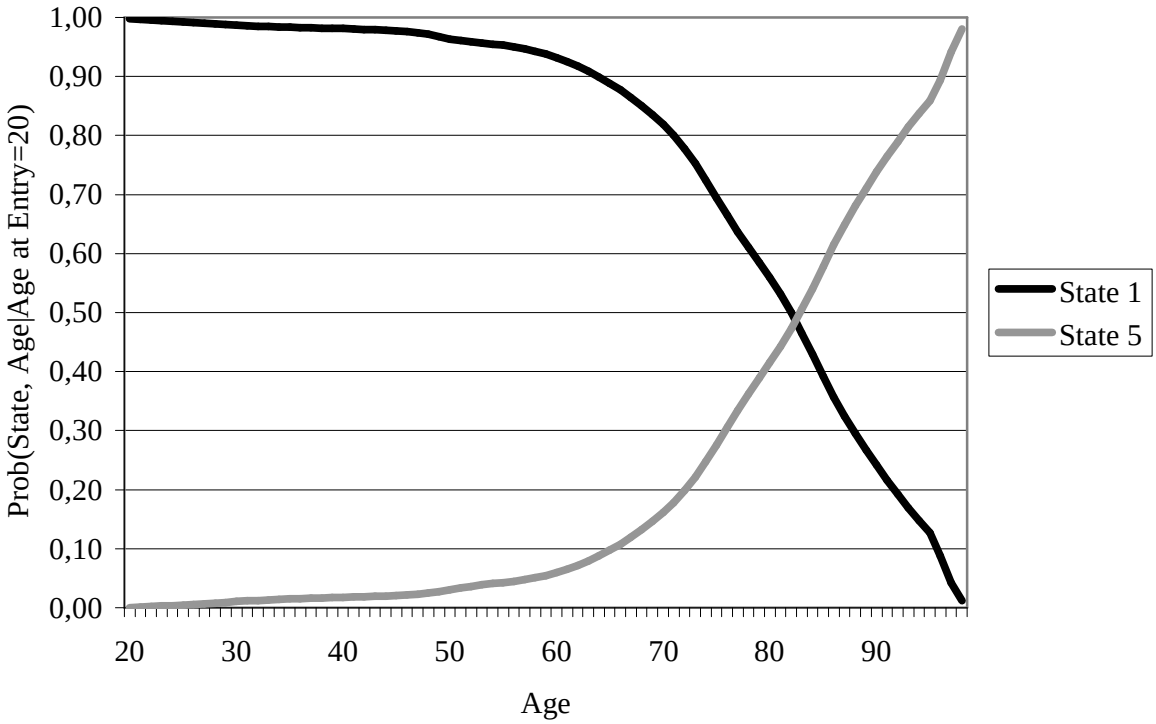

Figure 2.3: Probability of Being in State 1 / State 5 (Males)
Probability conditional on Being in State 1 at the Age of 20

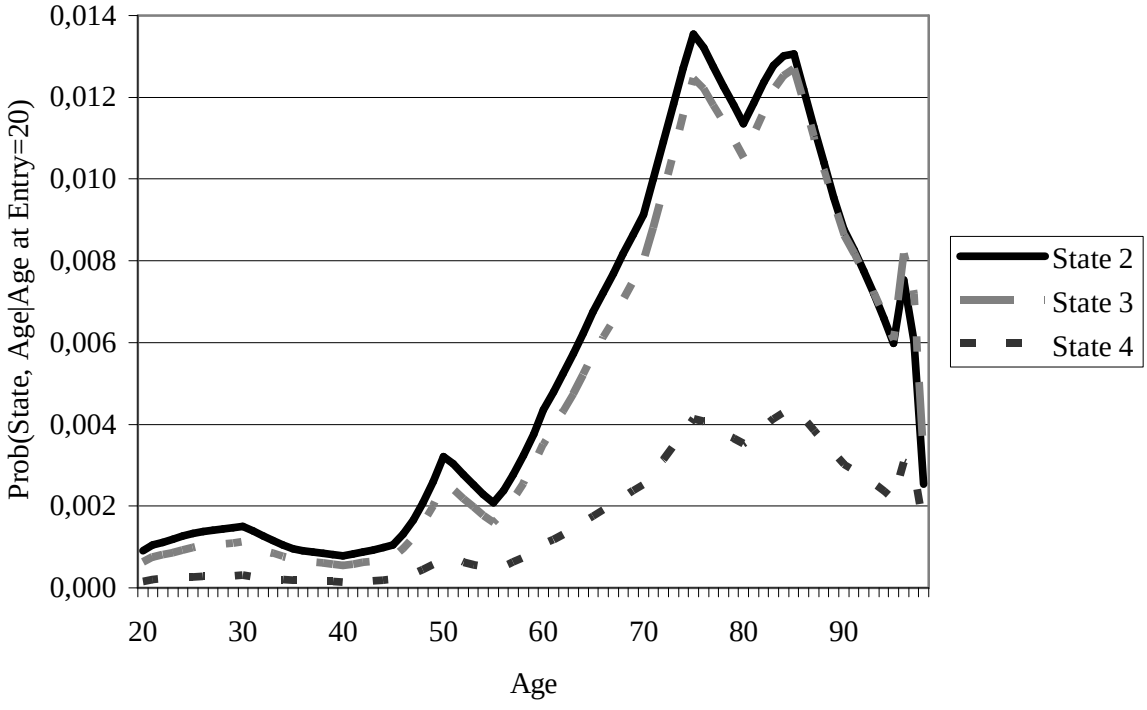

Figure 2.4: Probability of Being in State 2 / State 3 / State 4 (Males)
Probability conditional on Being in State 1 at the Age of 20
Note: the falling lines are due to persons who die.

probabilities from state one to the five states have to be considered. The result is a matrix of the dimension 79x5x79 for each gender. Figures 2.3-2.5 show how the transition probabilities change across life for j=0.[57]

In Figure 2.3 one can see that the conditional probability for a man to be healthy in a certain age (conditioned on being in perfect health at the age of 20) falls from 100% at the age of 20 to 1.2% at the age of 98% – and vice versa for the death probability. As shown in Figure 2.5 for living persons the probability of becoming a nursing case (again if the person is in perfect health at the age of 20) is rising constantly with age.

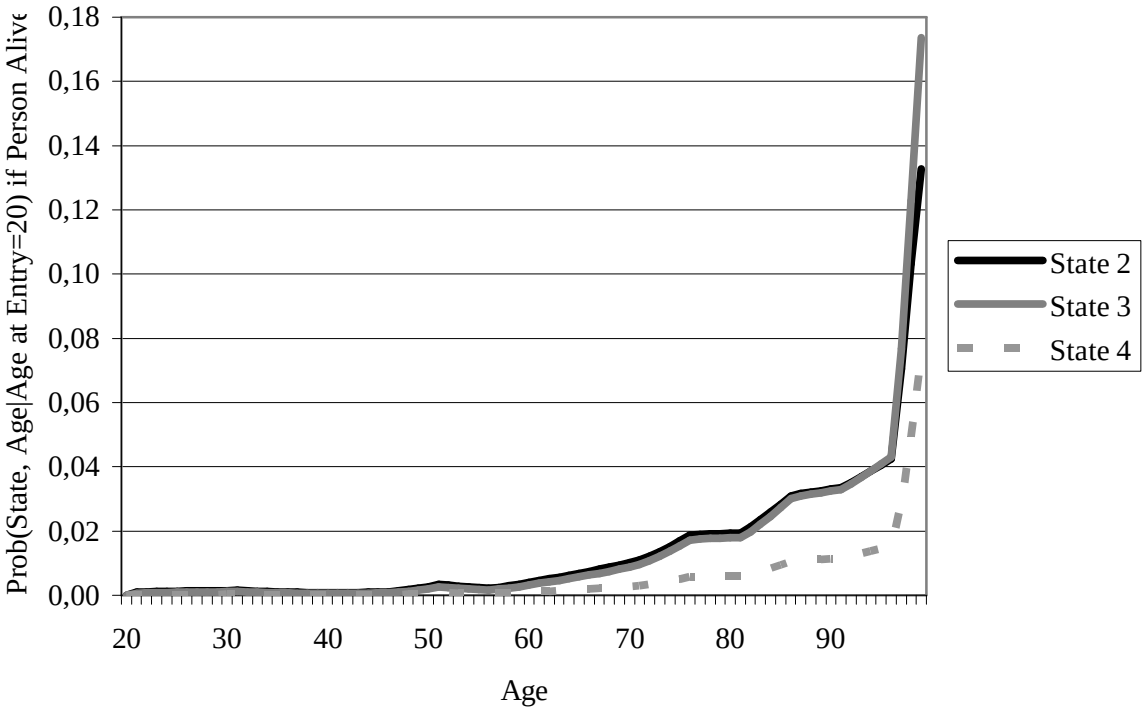

Figure 2.5: Probability of Being in State 2 / State 3 / State 4 if Person Alive (Males)
Note: These are conditional probabilities conditioned on "Being in state 1, 2, 3 or 4" → probabilities are higher here than in Figure 5. As state 5 is no longer in the Figure, the curves no longer fall within five-year-intervals.

2.5.2.3 Validation via "Stylized Facts"

One possible way to validate the transition probabilities is to compare them to probabilities which have already been published for Germany. I will refer to two sources: an article by Ziegler and Doblhammer (2005b) and the federal statistics of Germany (Statistisches Bundesamt (2007)).

Ziegler and Doblhammer (2005b) base their findings on data from the German Socio-Economic Panel. They point out that age is the most important risk factor for the transition into LTC. The risk rises exponentially with age. Furthermore, incidence is said to be the same

[57] Note that the spline curve characteristic is due to the definition of five year age splines.

for men and women, but prevalence is much higher for women. This means that women, once they entered LTC, stay frail for a longer period than men given that mortality rates are much lower for women in all ages.
Federal statistics as Statistisches Bundesamt (2007) confirm the facts about the exponential influence of age and the high prevalence of women in LTC. The difference between genders concerning the share of the population in LTC is increasing with age, too. While about 5% of both sexes are in LTC in the group of 70-75 year olds, 69% of female over 95 year olds and only 29% of males over 95 years are in LTC.

My transition probabilities suggest the same findings. The LTC risk increases strongly with age: 1.33% (0.79%) of 65 year old men (women) in good health is in LTC the following year. For 95-year-olds, probabilities reach 8.35% (9.89%). Concerning mortality rates, my probabilities show values that are consistent with Ziegler and Doblhammer (2005b): 26.36% of men in care level 1 at the age of 65 die within the following year, whereas it's only 10.17% for women. 10 years later, 44.16% of those men are dead while only 21.95% of women who were in care level 1 at the age of 65 have died.
A similar mortality difference can be observed for the healthy 95-year-olds: 6.99% of the men in this group die within the next year as compared to 1.14% of women.
On the one hand, the fact that women stay in LTC longer than men can be proved by lower mortality rates. On the other hand, probabilities even differ for the 65-olds: 31.16% (only 16.58%) of women (men) in care level are still in LTC the following year. This effect is much stronger for 95-year-olds, where 38.71% of the women and only 4.91% of the men in care level one are still in LTC one year later.
This comparison shows that my transition probabilities are in order with existing research about LTC risk, mortality rates and their difference for both sexes. Thus, this is one way to successfully validate the estimated probabilities.

2.5.3 A hedonic Regression of Insurance Premiums in the German Market for SLTC

In order to calculate a realistic mean premium I use the results of a hedonic regression I ran with the offers of 24 German insurers for daily allowance SLTC insurance.
The dataset includes 24 insurers with, in total, 35 insurance plans. The price for an insurance policy basically depends on the customer's age and gender. Furthermore, contract-specific components influence the price, such as the benefits in each care level as shares of the benefit payment in care level three (e.g. 25% in level one, 50% in level two and 100% in level three),

the maximum age at entry and the time after buying the insurance that the customer has to wait until he can receive benefits (usually up to 3 years).

Benefit payments are usually offered as multiples of a 5€ payment in care level three. There is a base price for a 5€ benefit unit. The total monthly premium is the combination of the base price and the according factor to get the contracted benefit amount. E.g., if a 5€ daily allowance costs 0.50€ per month, a 20€ daily allowance costs 2€ per month. Premiums usually stay constant for a life-time[58] unless optional arrangements to increase coverage and thus premiums automatically after a certain period have been agreed on.

Table 2.7 gives an overview over the included insurance companies and their insurance plans.

Insurer	Plan 1	Plan 2	Plan 3
Allianz Private Krankenversicherungs-AG	PZT		
Alte Oldenburger Krankenversicherung	PT		
ARAG Krankenversicherungs-AG	Tarif 69		
Barmenia Krankenversicherung a.G.	PT1	PT3	
Bayerische Beamtenkrankenkasse	PflegeOPTIMAL	PflegeKOMPAKT	
CENTRAL Krankenversicherung	EPTN1	EPTN2	
DBV-Winterthur Krankenversicherung	PTG 3	PTG DYN	
Debeka Krankenversicherungsverein	PVZ		
Deutscher Ring	PTG 1	PTG 2	
DEVK Krankenversicherungs-AG	PT/B	PT/B3	
DKV Deutsche Krankenversicherung	PET		
Gothaer Krankenversicherung	PT		
HanseMerkur Krankenversicherung	PTA		
HUK-Coburg-Krankenversicherung	PT	PT3	
INTER Krankenversicherung	PTN		
LVM Krankenversicherungs-AG	PZT		
Münchener Verein	Tarif 420	Tarif 423	
Nürnberger Krankenversicherung	PT		
PAX-Familienfürsorge Krankenversicherung	P EU		
R+V Krankenversicherung	PT	PT3	
SIGNAL Krankenversicherung	EPT		
Süddeutsche Krankenversicherung	PE1	PE2	PE3
uniVersa Krankenversicherung	PTK		
VICTORIA Krankenversicherung	PZ		

Table 2.7: Insurers and Insurance Plans

Variables	Mean	Std	Label
accidentwait	0,56	0,50	1 if standard waiting time does not apply in case of an accident
age	41,35	16,33	age at the moment of buying the insurance
age_gender	20,67	23,68	Interaction of age and gender: equals the age for men and zero for women
alwayswait	0,00	0,00	1 if there is a standard 3 year waiting time
level1_euro	0,76	0,82	€ paid in care level 1
level2_euro	2,28	1,61	€ paid in care level 2
level3_dummy	0,30	0,46	1 if only care level 3 triggers payments
lprice	0,69	1,05	logarithm of the price for a premium with 5€ daily allowance in care level 3
max_age	67,21	8,30	maximal entrance age to contract the insurance
neverwait	0,32	0,47	1 if there is no waiting time between contracting and receiving benefits
older	38,90	20,49	age if age>=22, 0 else
gender	0,50	0,50	1 if male, 0 if female

Table 2.8: Variables and Values (3676 Observations)

[58] See general assumptions in section 2.

Table 2.8 shows the variables in the regression and their descriptive statistics. The regression results are reported in Table 2.9. With this regression 92% of the price variation can be explained. For more details and a discussion of the results see Scharfenberg (2008).[59]

	(1)
gender	-0.227***
	(0.043)
age_gender	-0.00474***
	(0.00092)
level1_euro	0.176***
	(0.045)
level2_euro	0.173***
	(0.036)
level3_dummy	*-0.277***
	(0.13)
accidentwait	0.0219
	(0.068)
neverwait	-0.0132
	(0.071)
max_age	0.00970***
	(0.0026)
age	0.0672***
	(0.010)
age_sq	-0.000161
	(0.00012)
Constant	-3.083***
	(0.33)
Observations	3,676
R-squared	0.92

Table 2.9: Hedonic Regression Results
(standard errors in brackets, ***/**/* for significance at 1%/5%/10% level)

2.5.4 Estimation and Results

The payments P_S stay constant for different ages and ages at entry. As payments are linear in prices only the payment-price-ratio is important here. In order to facilitate the calculation a coverage of 40€ a day is assumed. As shown in Table 2.10, this would be just enough to cover the difference between the arising costs in care level three and the LTC payments.

	Care Level		
	1	**2**	**3**
Costs (nursing expenses + board & lodging)	1,854	2,280	2,706
Payments out of LTC insurance	1,023	1,279	1,432
Difference	831	1,001	1,226
SLTC insurance payments (40€/day, average payments in level 1 and level 2: 15%, 45%)	180	540	1,200

Table 2.10: Care Costs and LTC Payments[60]

The insurance prices are calculated using the hedonic function of chapter 2.2.3. To be more realistic the function is calculated step-wise in five groups taking into account the different

[59] In Scharfenberg (2008), I use some more variables in a more recent regression. However, as these variables are not significant and the results do not change substantially, the original values are being used here.
[60] Statistisches Bundesamt (2007), p. 15; Bundesministerium für Gesundheit (2007a), p. 4.

maximum ages at entry and then smoothed via an additional regression using the predicted prices from the step-wise function as the dependent variable and age and squared age as independent variables.[61] Figure 2.6 shows the resulting price function as well as the step-wise function.

Premiums stay constant for individuals that contract the insurance after the age of 21. Before that they pay a special rate that is adapted to the adult rate after another risk examination at the age of 21. This is why individuals below the age of 22 are excluded from the analysis.

Estimations for the whole market, i.e. based on the hedonic price function, show that loads are very high in the German market. See Figure 7 for a graphic analysis. For the estimation with a 4% discount rate, female loads start at 0.27, loads for men at 0.05 at the age of 22. In the case of a 6% discount rate, the load factors are higher as future benefits – which will be paid many years after buying the insurance – are discounted by a higher factor and thus lower in their EPDV than the premiums that start from the first period.

An analysis of single insurers (as shown in Figure 2.8) shows that the difference between the loads for men and women approaches zero with rising age. As insurers do not have an incentive to discriminate one gender, these equalized loads show that the model replicates the real process of calculating insurance prices to a certain extent.

Figure 2.6: Annual Premiums (Coverage: 40€ / Day)
Note: Premiums for Males; for Females similar

[61] The groups are divided according to the maximum ages at entry 60, 65, 70, 75 and 100.

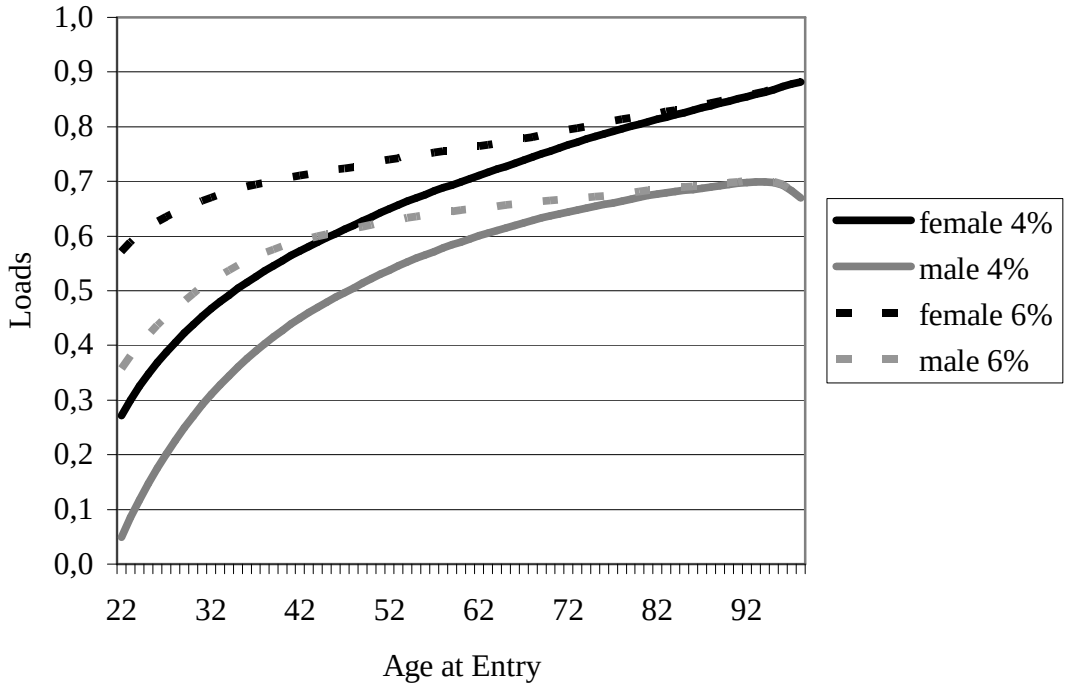

Figure 2.7: Loads
Female Loads and Male Loads with 4% and 6% discount rate

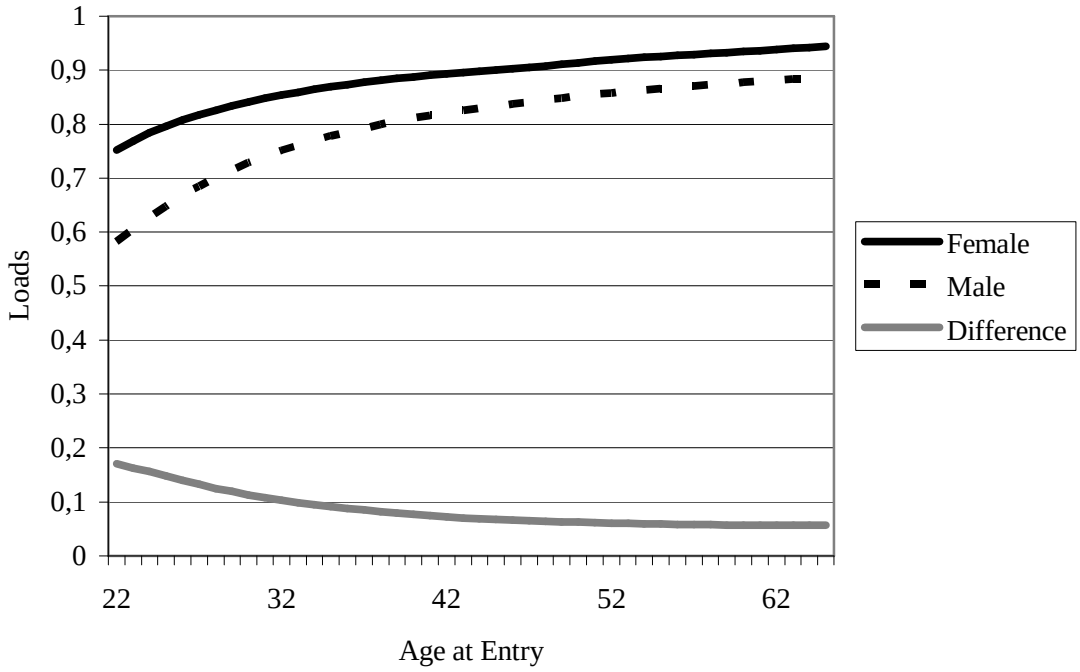

Figure 2.8: Loads for a Single Insurer
6% discount rate

2.5.5 Unconditional Probabilities and Claims Amounts per Risk

German insurers do not calculate their SLTC premiums based on transition probabilities into care levels. They use the claims amount per risk, i.e. the average benefit payment for an indi-

vidual of a certain age and gender. The amount is deducted from actual benefit payments which are not published.[62] However, these amounts do not take the health state of the benefit receivers at the age at entry into account and do not vary for different ages at entry. If a 20 year old man and a 65 year old man enter the insurance, the insurer assumes the same probability for both to be a nursing case at the age of 70. In the transition probability model, LTC probabilities are much lower for the 65 year old as they are conditioned on him being in perfect health at the age of 65. As a consequence only claims amounts at the age of 70 per 65 year old healthy man are taken into consideration – and not those that include men who are already in LTC at the age of 65. With the concept of unconditional probabilities, I try to replicate the idea behind the insurers' approach without needing the data on claims amounts per risk.

The unconditional probability for a person to be in a health state s' in time t is given by $P(s_t = s_t')$ whereas the conditional transition probability is given by

$$P(s_t = s_t' \mid s_{t-1} = s_{t-1}'), \tag{2}$$

taking the health state in the previous year into consideration. Hence the unconditional probability is the sum of the conditional probabilities over all states:

$$P(s_t = s_t') = \sum_{s_{t-1}'=1}^{5} P(s_t = s_t' \mid s_{t-1} = s_{t-1}'). \tag{3}$$

In t-n the insurance only allowed persons who were in health state one to enter. Subsequently the transition probability model calculates the probability $P(s_t = s_t' \mid s_{t-n} = 1)$ while the unconditional probability remains $P(s_t = s_t')$. Therefore,

$$P(s_t = s_t') \geq P(s_t = s_t' \mid s_{t-n} = 1) \tag{4}$$

$$\text{or, if } \sum_{s_{t-n}'=2}^{5} P(s_t = s_t' \mid s_{t-n} = s_{t-n}') > 0, \tag{5}$$

$$P(s_t = s_t') > P(s_t = s_t' \mid s_{t-n} = s_{t-n}'). \tag{6}$$

Therefore LTC probabilities are lower in the transition probability case as only the subset of healthy persons in t-n is used to calculate probabilities. But as LTC probabilities are higher in the unconditional model, load factors have to be lower. This is due to the fact that benefits are weighted with a higher probability than in the transition probability model.

[62] Email from Isabella Osterbrink (PKV), 08-09-07.

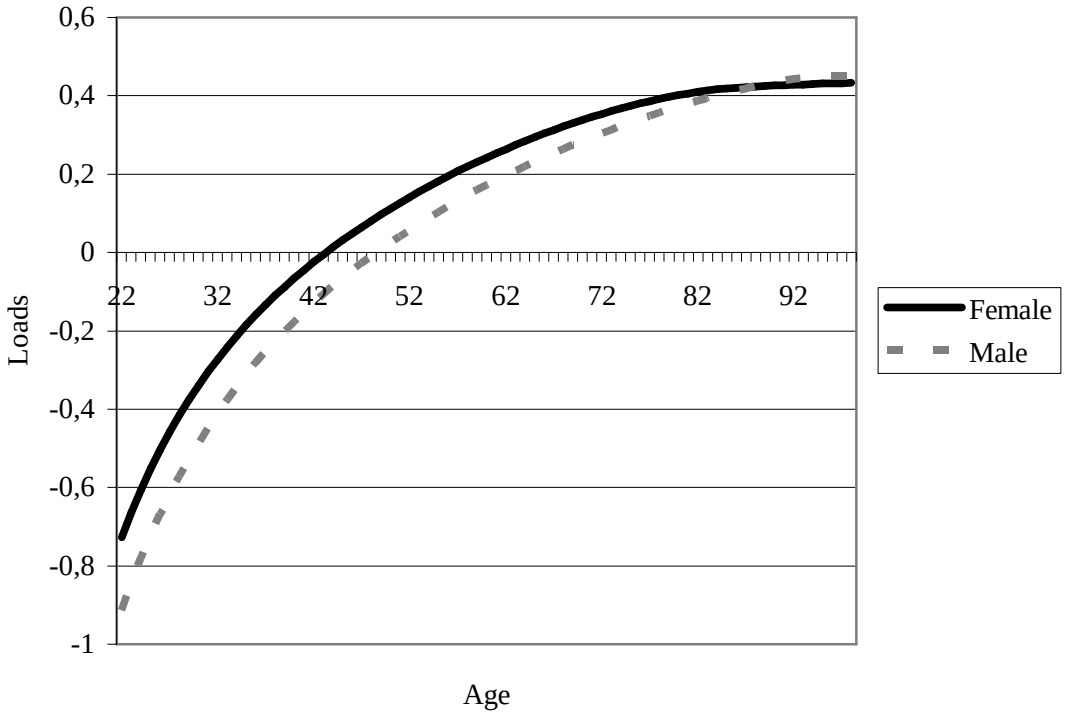

Figure 2.9: Loads calculated with Static Probabilities
6% discount rate

Figure 2.9 shows that loads are indeed much lower for the unconditional case. They now range from -0.91 (-0.73) for men (women) at the age of 22 to 0.45 (0.43) at the age of 98.
Even if according to this unconditional approach load factors are much lower, it is important to keep in mind that the load model is based on transition probabilities and that this approach has only been used to demonstrate the difference between transition probabilities and unconditional probabilities as employed by German insurers.

2.5.6 Further Applications of the Model

With the EPDV model, a comparison of different contracts with varying terms is possible. On the one hand, insurance companies can be compared in order to choose the optimal company. On the other hand, different contract models within a company can be compared.
An example for the latter is shown in Figure 2.10. Plan 1 offers 100% benefit payment limited to care level three and allows for an entrance age up to 75 years. In levels one and two the insured still has to pay insurance premiums. In plan 2 the age at entry may not exceed 65 years. Benefit payments reach 40% in level one, 70% in level two and 100% in level three. As soon as the insured enters the first care level, his monthly premium is set to zero.

Plan 1 is far less expensive than plan 2. However, the load calculation shows that in expectation plan 2 pays more. Due to the fact that care level three is reached less often than level one and level two, most insured will never receive benefits from plan 1.

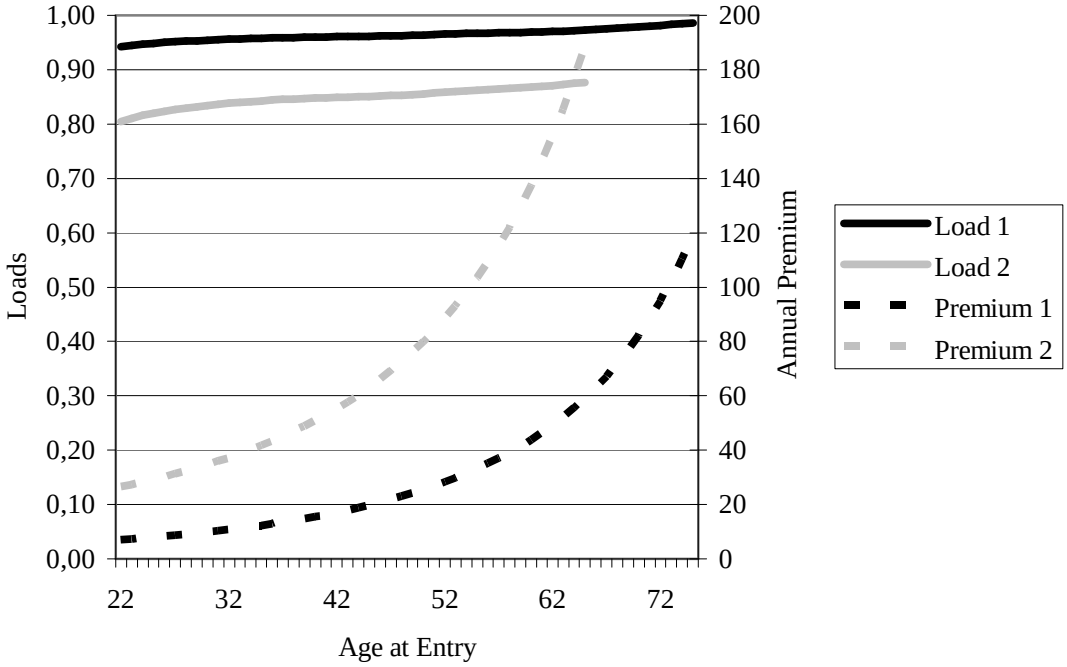

Figure 2.10: Contract Comparison (Female)
6% discount rate, Annual Premium (Coverage: 40€ / Day)
Contract 1 only pays in level 3; Contract 2 pays 40%-70%-100% in Care Levels 1-2-3

2.6 Market Failure as an Explanation for high Loads

Brown and Finkelstein (2007) define four sources of supply side failures in the market that may lead to elevated prices, i.e. high load factors, or quantity rationing. This sections' intention is to identify possible types of market failures as reasons for high loads.

2.6.1 Asymmetric Information

According to Brown and Finkelstein (2007) high loads may have their origin in asymmetric information. As a successful risk examination is the basic condition to enter SLTC insurance adverse selection is not very likely to be based on unidentified health risk factors. However, more detailed information on health is needed to be able to rule out adverse selection based on risk type.

Finkelstein and McGarry (2004) point to another source of asymmetric information: private information about one's risk preferences. They argue that those whose preference for insurance is above the average which the insurer observes have below average use of nursing homes as this group consists of more cautious and wealthier individuals or persons who are

very likely to invest in health activities that the insurer cannot observe. As no detailed information is available about the bought policies and the insured benefits a test for asymmetric information as proposed by Finkelstein and McGarry (2004) or Finkelstein and Poterba (2006) cannot be performed here. However, as the rich are more likely to buy SLTC insurance (see chapter 3 for the Probit Regression) and as rich persons tend to be healthier than the average[63] there seems to be risk preference based adverse selection in the market.

For the United States Finkelstein, McGarry and Sufi (2005) prove that persons who lapse an insurance contract are found to be one-third less likely to subsequently be in a nursing home than those who keep their contract. The Mikrozensus Panel 1996-1999 only contains four subsequent periods; therefore later nursing home use cannot be taken into account in my analysis. Based on the available data only preliminary evidence can be found on the topic of dynamic inefficiencies; in χ^2 tests between persons who drop out of SLTC insurance from 1996-1997 and covariates only few correlations can only be found: those who have kids drop out more, the poorest drop out less, as do the richest. No other correlation, e.g. with education, sex and other income levels, can be proven. This makes dynamic inefficiencies very unlikely to happen in this context as no concrete risk group is dropping out of the insurance. Especially there is no proof of the healthy and wealthy dropping out. Furthermore, as there is no question in the Mikrozensus Panel 1996-1999 about whether someone dropped out of SLTC insurance I cannot distinguish between lapses and data problems. However, there is evidence that people's expectations about their own nursing home use are quite accurate.[64] Therefore it is possible that those with a very low risk drop out of the insurance over time as they can concretize their expectations over the years.

An analysis of moral hazard would only be possible with a data set which tracks customers after buying the insurance as the insured population should have a higher probability to become a nursing case after contracting the insurance than the uninsured population.

2.6.2 Transaction Costs

Transaction costs may have their origin in insurer's costs. Insurance companies face administrative costs that have to be covered by the premiums. Figures for the whole market of private health insurers indicate that the administrative costs are about 2.85%, acquisition costs are 8.61% of gross premiums in total.[65] Therefore 11.4 Cent of each paid Euro is to cover administrative and acquisition costs. Furthermore, a safety loading of at least 5% has to be added to

[63] See, e.g. Smith (1999) for an analysis of the correlation between health and wealth.
[64] See Kleinjans and Lee (2006) for an overview on the subject and a survey on existing literature about nursing home expectations.
[65] PKV (2006), p.97.

the premium.[66] Unfortunately, figures about the cost of claims settlement are not available. I estimate these costs at another 7% after speaking with an actuary.
This is a total of 23.5% for insurer costs meaning that 23.5 Cent of each premium Euro is used to cover costs. However, even after deducting these costs, loads stay very high.

2.6.3 Imperfect Competition

The effect of transaction costs can be aggravated if competition in the market is not working. Prices actually vary a lot in the market – even for identical contracts – and the market leader is among the most expensive insurers in place. Clearly, competition in the market is not perfect. See Scharfenberg (2008) for a detailed characterization of the SLTC insurance market.

2.6.4 Uninsured Aggregate Risk

In addition to facing an individual risk of becoming a nursing case, the whole group of insured persons is exposed to the intertemporal risk of rising LTC costs. If the insurer tries to cover this risk by an additional premium prices can be elevated. Daily allowance insurance policies avoid this risk by fixing benefits to a certain amount of money which is independent of actual LTC costs. To cover inflation and rising costs, most contracts include an option to raise premiums and benefits over the years. However, this uninsured aggregate risk cannot lead to elevated prices in this market.

2.7 Discussion

This is the first time probabilities for the transition into LTC are calculated on a large data basis. Brown and Finkelstein (2007) use a dynamic continuous-time Markov chain model developed by James Robinson (1996) to calculate the transition probabilities in the first transition matrix. In this study a simpler approach is followed. The first matrices are calculated using an Ordered Probit model. Then the Markov property is assumed. Of course, if this property is not fulfilled, the more complicated dynamic approach will be the better option. However, the resulting transition probabilities are consistent and the probability curves are shaped as expected. An actuary appreciated my Markov solution as an appropriate solution.

In Germany, premiums are calculated according to insurance law (Kalkulationsverordnung). This leads to a different calculation than the one presented here. To assure a certain security level, the expected average probabilities or claims amounts per risk have to be taken as a

[66] §12 c), VAG (2007).

lower bound for the probabilities in the premium calculation. This leads to higher probabilities in the calculation – and thus to higher premiums. Even taking into account the insurance costs, premiums will still have to be higher to follow insurance law. See Scharfenberg (2008) for an overview on general calculation rules for premiums in the German health insurance market.
Furthermore, insurers have to calculate with lapse probabilities. A certain percentage of the insured cancel their contracts before they receive any benefits. This reduces the payments for the insurer and decreases the premium. As the Mikrozensus Panel does not offer specific data about lapses, new data sources would be needed to account for this possibility.

People's individual decisions strongly depend on their level of risk aversion. The "fair load of zero" defined by the EPDV model is a criteria for risk neutral individuals only. Very risk adverse persons will find much higher load factors acceptable. Unfortunately, no information on risk aversion is available in the Mikrozensus Panel 1996-1999. Moreover, people might not have the computational skills to calculate the loads on a certain model. They would not know that the cheaper contract that only pays benefits in care level three has a lower EPDV than the more expensive contract. The fact that men buy substantially more SLTC insurance policies while the load factors barely differ between sexes underlines this point.
Finally, the common way to sell SLTC insurance policies is to sell them in a bundle with retirement provisions or health insurance.[67] It can be assumed that people do not compare contracts that are offered by different insurers; they rather just buy the insurance from their usual provider. The consequence is that the price does not have a very wide influence on the buying decision.

2.8 Conclusion

This study shows that factors on the supply side of the German market for SLTC insurance limit demand in the market. According to the EPDV model by Brown and Finkelstein (2004a) prices in the market are too high. This can be caused by supply side failures in the market such as asymmetric information or imperfect competition.
To calculate the EPDV model, I estimate probabilities for the transition between health, three care levels and death. The Mikrozensus Panel 1996-1999 is the first data set to offer the possibility to estimate these probabilities on a large number of observations. After predicting

[67] See, e.g., Allianz (2007).

probabilities for one year transitions based on the estimates of an Ordered Probit model I use Markov chains to estimate probabilities for later years. In order to compute the expected values, I use actual benefit payments and premiums that are currently available in the SLTC insurance market. Values are discounted by factors of 4% and 6%.
The resulting load factors show that according to the EPDV model premiums in the market as a whole are strictly increasing with the age at entry into the insurance and that they are much higher than zero and approaching one for high ages. High load factors can hint to supply side failures in the market. While imperfect competition and high transaction costs are very likely to be a problem, the coverage of the risk of rising LTC costs can be ruled out as an origin to high loads as daily allowance insurance policies are independent of the actual LTC costs. More data on the insured population is needed in order to identify problems of adverse selection and moral hazard in the market.

Another application of the load model is a comparison of different insurance contracts. It can be shown that within an insurance company, cheaper contracts with fewer benefits can actually have higher load factors than more expensive contracts. Therefore, one possible extension of the model is to develop a comparison tool to help customers to choose the best available contract in the market.

Finally, supply side failures only account for some of the possible factors that limit the SLTC insurance market size. Problems on the demand side seem to be of significant influence on the market, too. They include limited rationality of customers, especially misconceptions about the need and the possibility to insure the LTC risk on an individual basis.

References

Allianz AG (2007): Der Allianz Schutzbrief für alle ab 55. http://www.allianz.de/produkte/personen/schutzbrief_55plus/steckbrief/index.html (accessed in December 2007).

Brown, Jeffrey R. and Amy Finkelstein (2004a): Supply or Demand: Why is the Market for Long-Term Care Insurance So Small? NBER Working Paper 10782.

Brown, Jeffrey R. and Amy Finkelstein (2004b): The Interaction of Public and Private Insurance: Medicaid and the Long-Term Care Insurance Market. NBER Working Paper 10989.

Brown, Jeffrey R. and Amy Finkelstein (2006): The Interaction of Public and Private Insurance: Medicaid and the Long-Term Care Insurance Market. Revised and resubmitted, *American Economic Review*.

Brown, Jeffrey R. and Amy Finkelstein (2007): Why is the Market for Long-Term Care Insurance So Small? *Journal of Public Economics*, forthcoming.

Brown, Jeffrey R., Norma B. Coe and Amy Finkelstein (2006): Medicaid Crowd-Out of Private Long-Term Care Insurance Demand: Evidence from the Health and Retirement Survey. NBER Working Paper 12536.

Bundesanstalt für Finanzleistungsaufsicht (2007): Gesetz über die Beaufsichtigung der Versicherungsunternehmen (Versicherungsaufsichtsgesetz - VAG). http://www.bafin.de/gesetze/vag.htm (accessed in December 2007).

Bundesministerium für Gesundheit (2006): BMG-P-G500, Pflegeversicherung. http://www.bmg.bund.de/cln_040/nn_603392/SharedDocs/Publikationen/Pflege/g-500,templateId=raw,property=publicationFile.pdf/g-500.pdf (accessed in December 2007).

Bundesministerium für Gesundheit (2007a): Zahlen und Fakten zur Pflegeversicherung. http://www.bmg.bund.de/cln_040/nn_773096/SharedDocs/Download/DE/Themenschwerpunkte/Pflegeversicherung/Informationen/ZahlenFakten,templateId=raw,property=publicationFile.pdf/ZahlenFakten.pdf (accessed in December 2007).

Bundesministerium für Gesundheit (2007b): Das bringt die Reform der Pflegeversicherung. http://www.bmg.bund.de/cln_040/nn_604244/DE/Themenschwerpunkte/Pflegeversicherung/Pflegereform-17-10,param=Links.html#doc1173806bodyText2 (accessed in December 2007).

Bundesministerium für Gesundheit (2007c): Pressemitteilung Pflege. http://www.bmg.bund.de/cln_040/nn_604244/DE/Presse/Pressemitteilungen/Presse-4-2007/pm-17-10-07.html (accessed in December 2007).

Czado, Claudia and Susanne Gschlößl (2002): Modeling of transition intensities and probabilities in a German long term care portfolio with known diagnosis. Sonderforschungsbereich 386, Paper 302 (2002).

Finanztest (2006): Pflege für alle. *Finanztest*, (4), 68-78.

Finkelstein, Amy and James Poterba (2006): Testing for Adverse Selection with ‚Unused Observables'. NBER Working Paper 12112.

Finkelstein, Amy and Kathleen McGarry (2006): Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market. *American Economic Review*, 96(4), 938-958.

Finkelstein, Amy, Kathleen McGarry and Amir Sufi (2005): Dynamic Inefficiencies in Insurance Markets: Evidence from Long-Term Care Insurance. *American Economic Review Papers and Proceedings*, (95), 224-228.

Häcker, Jasmin and Bernd Raffelhüschen (2004): Denn sie wussten, was sie taten: Zur Reform der Sozialen Pflegeversicherung. *Vierteljahrshefte zur Wirtschaftsforschung* (73) (2004), 158-174.

Helms, Florian, Claudia Czado and Susanne Gschlößl (2004): Calculation of LTC Premiums based on direct estimates of transition probabilities. Sonderforschungsbereich 386, Paper 393 (2004).

Kleinjans, Krisin J. and Jinkook Lee (2006): The link between individual expectations and savings: Do nursing home expectations matter? Department of Economics, Working Papers 2006-05, Department of Economics, University of Aarhus.

Kunreuther, Howard et al. (1978): "Disaster Insurance Protection: Public Policy Lessons." New York: Wiley.

Milbrodt, Harmut (2005): Aktuarielle Methoden der deutschen Privaten Krankenversicherung. DGVFM: *Schriftenreihe angewandte Versicherungsmathematik*, 34.

Pauly, Mark (1990): "The Rational Nonpurchase of Long-Term-Care Insurance." *Journal of PoliticalEconomy*, 98(1), 153-168.

Robinson, James (1996): A Long-Term Care Status Transition Model. Presented at the 1996 Bowles Symposium on the Old-Age Crisis: Actuarial Opportunities. Georgia State University, September.

Rothschild, Michael and Joseph Stiglitz (1976): Equilibrium in Competitive Insurance Market: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics*, 90(4), 629-649.

Scharfenberg, Veronika (2008): Private Supplementary Long-Term Care Insurance in Germany: Characterization of a Rising Market. Working Paper.

Schulz, Erika, Reiner Leidl and Hans-Helmut König (2004): The impact of ageing on hospital care and long-term care – the example of Germany. *Health Policy*, (67), 57-74.

Sloan, Frank and Edward Norton (1997): "Adverse Selection, Bequests, Crowding Out, and Private Demand for Insurance: Evidence from the Long-Term Care Insurance Market," *Journal of Risk and Uncertainty* (15), 201-219.

Smith, James P. (1999): Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status. *Journal of Economic Perspectives*, 13(2), 145-166.

Statistisches Bundesamt (2007): PFLEGESTATISTIK 2005: Pflege im Rahmen der Pflegeversicherung. 4. Bericht: Ländervergleich – Pflegeheime. https://www-ec.destatis.de/csp/shop/sfg/bpm.html.cms.cBroker.cls?CSPCHD=0010000100003ewgkgHc000000IoUvNZgyoaYXTYVEQErYMA--&cmspath=struktur,vollanzeige.csp&ID=1020079 (accessed in December 2007).

Statistisches Bundesamt (2007a): PFLEGESTATISTIK 2005: Pflege im Rahmen der Pflegeversicherung. Deutschlandergebnisse. https://www-ec.destatis.de/csp/shop/sfg/bpm.html.cms.cBroker.cls?cmspath=struktur,vollanzeige.csp&ID=1019863 (accessed in March 2008).

Verband der privaten Krankenversicherung (PKV) (2000): Die private Krankenversicherung 1999/2000. Zahlenbericht. Köln 2000 http://www.pkv.de/downloads/Zb00.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2001): Die private Krankenversicherung 2000/2001. Zahlenbericht. Köln 2001 http://www.pkv.de/downloads/Zb01.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2002): Die private Krankenversicherung 2001/2002. Zahlenbericht. Köln 2002 http://www.pkv.de/downloads/Zb02.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2003): Die private Krankenversicherung 2002/2003. Zahlenbericht. Köln 2003 http://www.pkv.de/downloads/Zb031.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2004): Die private Krankenversicherung 2003/2004. Zahlenbericht. Köln 2004 http://www.pkv.de/downloads/Zb04.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2005): Die private Krankenversicherung 2004/2005. Zahlenbericht. Köln 2005 http://www.pkv.de/downloads/Zahlenb_04_05.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2006): Die private Krankenversicherung 2005/2006. Zahlenbericht. Köln 2006. http://www.pkv.de/downloads/Zahlenb_05_06neu.pdf (accessed in December 2007).

Verband der privaten Krankenversicherung (PKV) (2007): Die Reform der Pflegeversicherung. Arbeitsversion. February. http://www.pkv.de/downloads/Bericht_Pflegeversicherung_22feb07_komplett.pdf (accessed in December 2007).

Ziegler, Uta and Gabriele Doblhammer (2005a): Reductions in the Incidence of Care Need in West and East Germany between 1991 and 2003: Compression-of-Morbidity or Policy Effect? Working Paper for IUSSP in Tours, France, July.

Ziegler, Uta and Gabriele Doblhammer (2005b): Steigende Lebenserwartung geht mit besserer Gesundheit einher. *Demographische Forschung aus Erster Hand*, 2(1), 1-2.

Zweifel, Peter (2005): The economics of long-term care: a survey. Contribution to the European Conference on Long-Term Care, Mannheim, October 21-22, 2005. ftp://ftp.zew.de/pub/zew-docs/veranstaltungen/LTC_Conference/25.a.pdf (accessed in December 2007).

3 Who becomes Self-Employed?
Evidence from the Mikrozensus Panel

3.1 Introduction

New firm formation is the basis for growth and competitiveness of an economy. While innovative new firms can enter new markets, all new enterprises can help create jobs and consolidate social security systems.

859,000 persons between 18 and 64 became self-employed as primary or secondary occupation in Germany in 2007. This is the lowest number since 2000. The gross effect on employment per new enterprise was at 1.9 employees per self-employed as a primary occupation and 0.3 employees per self-employed as a secondary occupation. In total, 455.000 full-time jobs were created by new self-employed persons in 2007. As much as 17-18% of all new self-employed were previously unemployed.[68]

Starting in 1986 the German government created a number of incentives to support the individual decision to become self-employed: until 2006, new entrepreneurs coming from unemployment were entitled to a bridging allowance that covered 170% of their last unemployment benefit. From 2003 on, as part of the so-called "Hartz-Reformen", they were entitled to the "Existenzgruendungszuschuss für Ich-AG" (EGZ), a fixed monthly amount between 240€ and 600€, for up to three years.

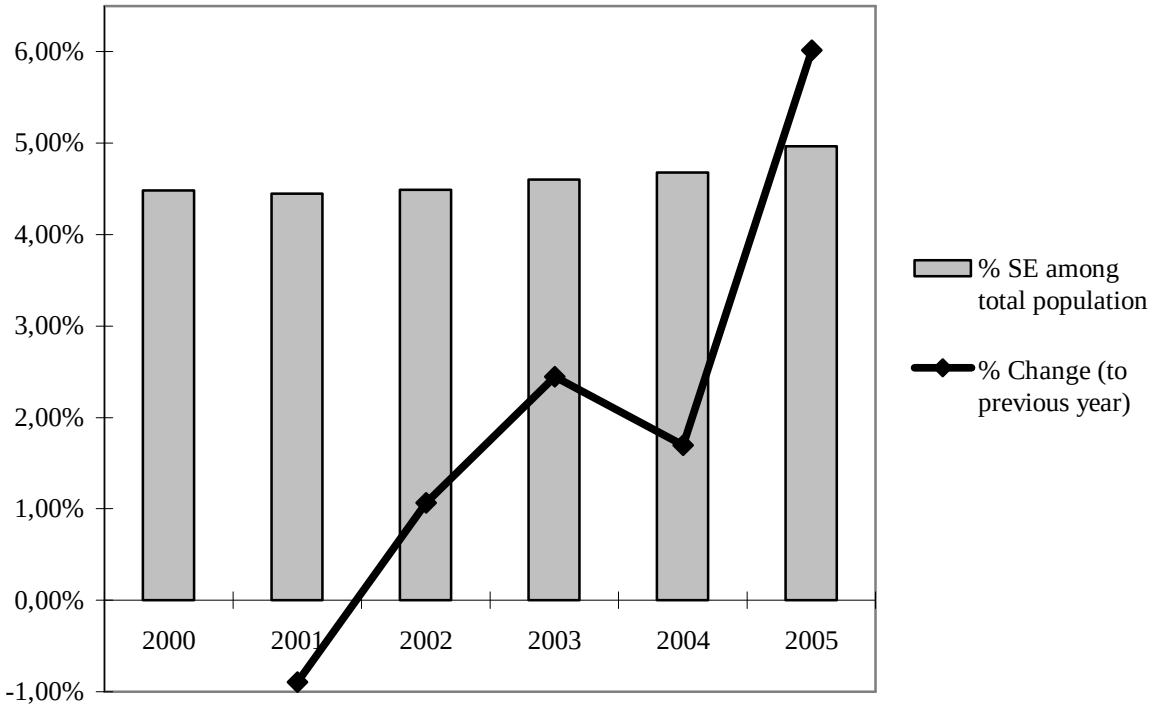

Figure 3.1: Share of Self-Employed among the Population
Based on Mikrozensus Cross Sections 2001-2005

[68] KfW (2008).

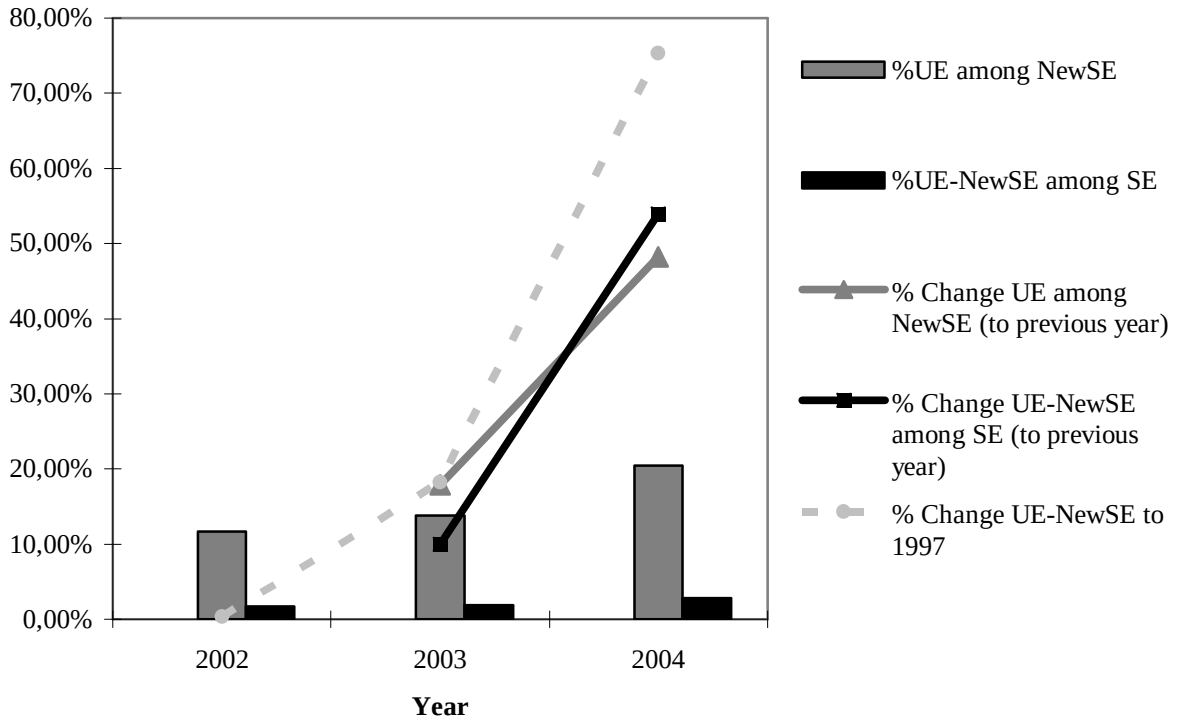

Figure 3.2: Share of previously Unemployed among Self-Employed / new Self-Employed
Shares of previously unemployed among all new self-employed, shares of previously unemployed new self-employed among all self-employed and change of the shares based on Mikrozensus Panel 2001-2004.
Change in the shares as compared to 1997 based on Mikrozensus Panel 1996-1999.

The fact that the share of previously unemployed persons among the new self-employed grew substantially between 2003 and 2004 suggests a success of this special incentive – especially considering that at the same time the total share of self-employed among the population grew slower than the year before. See figures 3.1 and 3.2 for the descriptive development of these shares.

In 2006, both support systems were replaced by the "Gruendungszuschuss", a subsidy for unemployed persons who want to become self-employed. It consists of a nine month basic support of the unemployment benefits plus 300€ and a potential six month follow-up support of 300€ per month. Besides these support systems which are designed to help unemployed individuals create their own business, there are a large number of subsidies and programs for employed persons, e.g. KfW StartGeld, KfW Unternehmerkredit, coaching programs etc.
The economic importance of new firm formation as well as the large amount of public spending on the above support systems cause large interest in identifying the characteristics of nascent entrepreneurs and the factors that influence the individual decision for self-employment.

Beside the possibility to evaluate existing political instruments this analysis can also help design incentives for future new entrepreneurs.
A true assessment of the pure influence of the introduction of the EGZ on the decision of the unemployed to become self-employed would only be possible with a comparable control group which did not have access to this incentive. As this control group is unavailable, it is hard to distinguish the effect of the incentive from other macroeconomic and microeconomic factors which might have influenced the decision. However, once the microeconomic variables of the decision are identified, the effect of the EGZ can be analyzed holding at least these factors constant.
Therefore, this paper has two main goals: first of all, I determine the influence of microeconomic variables like income, education and job status on the decision to become self-employed. Holding these variables' influence constant, I analyze the effect of the introduction of the EGZ.

This study is the first to use the Mikrozensus Panel 2001-2004 with about 700 transitions to self-employment per year. Prior studies for Germany are based on Mikrozensus cross sections (Pfeiffer 1999) or on the construction of pseudo-panels out of Mikrozensus cross sections as in Glocker and Steiner (2007). Generally, a lot of studies do not take unobserved individual effects such as entrepreneurial ability into account, famous examples being Evans and Leighton (1989) and (1990). This paper explicitly deals with unobserved effects. In order to do so I follow Chamberlain's Random Effects approach; in his model, the correlation of the individual effect with the regressors is represented by adding the individual means of the variables to the equation. It can be shown that results of hypothesis testing differ substantially from previous studies that did not account for unobserved heterogeneity.

My study results in several findings. Generally, unemployed and inactive persons are more likely to become self-employed than employed persons.[69] The longer an individual has been unemployed, the more likely she is to become self-employed. Besides, possessing liquid assets or having assets in the form of home ownership does not influence the decision to become self-employed. While personal net income is negatively correlated with that decision, the household net income per capita affects it in a positive way. Furthermore, a high level of education, i.e., having graduated from a university or from a university of applied sciences, is a negative contributor to the probability of becoming self-employed.

[69] For simplification I will refer to "employed" persons when speaking about those employed on the wage market as opposed to the self-employed.

Concerning the influence of the EGZ, unemployed persons are – ceteris paribus – significantly more likely to become self-employed in 2003 than in the years before: governmental incentives seem to work. Finally, a comparison with the classical random effects model shows large differences in coefficients: the influence of high education and personal net income on the decision to become self-employed is reversed in the classical model which considers the individual effects to be random.

The main policy implication is that the government's incentive mechanisms for the unemployed – especially the introduction of the EGZ – seem to work. However, the government should reconsider the economic contribution of those new firms: 41% of the previously unemployed new entrepreneurs in 2004 and 2005 had already left the market by 2007, partly because only one third of the new firms is created based on new ideas while two thirds are created out of dire straits. Therefore, employed and well educated persons should be encouraged to enter self-employment, given that their businesses tend to be more successful than those of the unemployed.[70]

In the next section, I describe the microeconomic background of the decision to become self-employed. While presenting the main hypotheses on the subject, I summarize the existing literature on self-employment.

In the third chapter I present the econometric model as well as a descriptive overview of the variables in the Mikrozensus Panel 2001-2004 which I use in my estimations.

Results for the entire sample, for the sample split by gender, and for the sub sample of the unemployed population are shown in section four. I also compare my findings to the results of a classical random effects Probit model and to the results of previous studies. Sections five and six discuss the results and give an overview of possible policy implications. Chapter seven concludes.

3.2 Microeconomic Background

Given basic conditions provided by the state, the individual decision to become self-employed depends on the utility of existing employment alternatives.[71] The most important factors which influence this utility are the material endowments to start a business, human resources, risk attitudes, the wish to be independent and the quality of social and familiar networks. If the utility of being self-employed is higher than the utility of the current job of an employed

[70] See KfW (2008).

[71] See Pfeiffer (1994) for a more detailed theory description.

worker or higher than the expected utility from employment of an unemployed person, she chooses self-employment.

3.2.1 Definitions

The definition of a self-employed person in this study follows the definition in the Mikrozensus: a self-employed individual is an economically active person who leads a business venture as an owner, co-owner or leaseholder on her own responsibility and without being bound to instructions and who bears responsibility for the development and the performance of the enterprise. The terms "entrepreneur" and "self-employed" are used synonymously in this paper because the data does not allow differentiating between, e.g., a self-employed artist and a start-up company.

Here, an unemployed person is defined according to the International Labour Organization (ILO) classification: someone who is jobless, looking for a job and who is available to the job market at short notice. This classification is used by the Mikrozensus opposed to the institutional classification according to which everyone who officially reported their status and everyone who is currently receiving unemployment benefits is defined as unemployed.[72]

3.2.2 Hypotheses and Literature Review

The main hypotheses concerning the effects of unemployment are called "push" and "pull" effect respectively. The push hypothesis argues that national or individual unemployment may force people into self-employment – or as the OECD (1986, p. 53) described it:

> *…in a slack labour market with few opportunities for paid employment, unemployed workers may seek self-employment as an alternative to joblessness, and multiple jobholders with secondary jobs in self-employment may lose their primary paid jobs, thereby becoming wholly self-employed.*

According to this hypothesis, self-employment should move counter-cyclically: whenever economic activity levels are low or falling, people tend to become self-employed.

The pull hypothesis, however, states that when economic activity levels are high or growing more people switch to self-employment and they are more successful with their ventures. Therefore self-employment should behave pro-cyclically.

Storey (1991) and Meager (1992) run surveys on previous studies on the subject, the latter proposing new approaches based on inflows and outflows from self-employment instead of the analysis of unemployment and self-employment stocks. For Canada Arai (1997) can only find a seasonal connection between unemployment and self-employment for the years 1961-

[72] See Schmidt (2000) for ILO classifications in the Mikrozensus.

1994 which can partly be explained by a large number of persons who run a business on a seasonal basis and register unemployed for the rest of the year.
While the push and pull hypotheses seem opposed to each other, they can be reconciled by looking at unemployment from different levels: individual unemployment is more likely to be a push factor while national unemployment or national economic activity is likely to be a pull factor. Ritsilä and Tervo (2002) run a nine year panel data analysis on Finnish Census Data to analyze the effects of unemployment on self-employment on three different levels: they find a positive influence of personal unemployment, a negative effect of national unemployment and no significant effect of regional unemployment on the decision to become self-employed.

Several macroeconomic studies focus on self-employment levels and unemployment rates, e.g. Glaeser (2007) who finds that demographic and industrial variation as well as area-level education account for a large part of the variation in self-employment rates across US cities. However, as this paper's focus is on the individual decision to become an entrepreneur, these aggregate studies are of minor interest here.

Another question regarding unemployment is whether the duration of unemployment influences the decision to become self-employed. While Ritsilä and Tervo (2002) find that with increasing duration of unemployment the propensity to start an own business is decreasing, Pfeiffer (1999) observes for Germany that the duration of unemployment only plays a role for salaried jobs, not for self-employment. He sees the reason in the fact that the entrepreneurs' customers do not know her employment history whereas a potential employer does.
In this study, the individual effect of unemployment and its duration will be analyzed.

The question whether liquidity constraints influence the individual decision to become an entrepreneur goes back to Joseph Schumpeter and Frank Knight. Schumpeter, e.g. Schumpeter (1946), sees the entrepreneur's role in identifying arbitrage opportunities in the economy which enables him to find a capitalist who will finance the enterprise and bear its risks. This means that a nascent entrepreneur does not have to bear any risk and is not bound by liquidity constraints if his idea is good enough. Knight (1921, 1964) argues that liquidity constraints do bind and that the entrepreneur has to bear most of his venture's risk The presence of asymmetric information and moral hazard problems in the capital markets can keep the entrepreneur from finding someone to finance his project. The result is that wealthier people are more likely to become entrepreneurs.

Evans and Jovanovic (1989) reject Schumpeter's hypothesis in favour of Knight's view as they find a strong relationship between individual wealth and the likelihood of becoming an entrepreneur. Petrova (2005) finds that part-time entrepreneurs do not seem to be financially constrained. Her intuition is that part-time entrepreneurs have their ventures in less capital intensive sectors.
An individual who thinks about becoming self-employed has to consider her opportunity costs. For employed workers, the opportunity cost of becoming self-employed is the wage. Therefore, a higher personal income can lead to a lower probability of the transition into self-employment. Evans and Jovanovic (1989) and Evans and Leighton (1989) confirm this hypothesis.
In my estimations I test for the effect of liquidity constraints in the decision process whether to become an entrepreneur.

Further studies treat the subject of risk aversion and entrepreneurship. Van Praag and Cramer (2001) analyze the subject in theory. Caliendo et al. (2006) estimate the influence of risk aversion on entrepreneurship and find that it only matters for employed individuals who decide to start a venture, not for the unemployed. However, as my data does not include variables on risk aversion, I cannot analyze it in my estimations.

For Germany, numerous studies concentrate on the characteristics of nascent entrepreneurs. Caliendo et al. (2006) rely on the German Socio-Economic Panel, Pfeiffer (1999) on a Mikrozensus cross section and Wagner (2002, 2007) on smaller data sets, i.e. the Regional Entrepreneurship Monitor Germany 2001 and 2003. They find low risk aversion, male gender and self-employed role models to be important factors for a person to become an entrepreneur. Constant and Zimmermann (2004) base their analysis on 19 waves of the German Socio-Economic Panel. They find transition probabilities from unemployment to self-employment to be much higher than those from employment. They partly explain this by the fact that for unemployed persons, self-employment can be shown to be an important channel back to regular employment. Köllinger and Schade (2005) point out that if Germans were as risk-loving, optimistic and self-confident as Americans, they would even found more enterprises per capita. In a pseudo panel analysis based on Mikrozensus cross sectional data, Glocker and Steiner (2007) find a positive influence of long-term unemployment on the self-employment rate. Besides, they find that men and persons between 35 and 40 years of age are most likely to become self-employed.

Concerning the EGZ, the most important study is the evaluation of the German Hartz programs by Caliendo, Steiner and Baumgartner (2007). They find that those who took part in the EGZ program are more likely to still be self-employed or employed after two and three years. However, as Eichhorst and Zimmermann (2007) point out, a large part of the sample was still receiving governmental subsidies at the time of the study. The actual effect of the EGZ on the dynamics of self-employment could not be analyzed in the literature because of the lack of a control group and suitable available data sets.[73] In my study, I can analyze the effect of the year 2003 on the unemployed, taking microeconomic variables into account.

The problem of an underlying, unknown entrepreneurial ability that influences the decision to become self-employed is often ignored in the literature. While van Praag and Cramer (2001) develop a theoretical model to deal with entrepreneurial ability, Evans and Jovanovic (1989) treat it as a part of a person's wealth. Unobserved heterogeneity is not considered in most of the studies that were mentioned above, e.g. Evans and Leighton (1989 and 1990). In Storey's 1991 survey of cross-sectional studies, none includes a hint to unobserved heterogeneity. Meager (1992) criticizes most of the previous studies but does not refer to a possible omitted variable bias. However, an omitted variable bias can lead to systematically over-/ underestimated coefficients and false test results. This paper will explicitly treat the problem of unobserved heterogeneity with the use of panel data methods.

3.3 Model and Data

In this study, I want to determine the factors which influence the individual's probability of becoming self-employed. This probability is a function of the person's individual characteristics such as age, gender and current job situation. In my regressions, after restricting the sample to those who are not currently self-employed, I use the dependent binary variable "NewSE" which is equal to one if someone is self-employed in the following period.

3.3.1 The Model

Model choice is guided by the plausibility of two assumptions: the first is strict exogeneity (1): conditioning on all explanatory variables $\underline{x}_i$ and the individual effect c_i, the expected mean of the error term u_{it} has to be zero. (1) has to hold in most panel data models in order to estimate consistent coefficients.

[73] See Kritikos and Kahle (2007), p. 52 for the problem of a suitable control group.

$$E(u_{it} \mid \underline{x}_i, c_i) = 0 \quad \forall t \tag{1}$$

$$E(c_i \mid \underline{x}_i) = E(c_i) = 0 \tag{2}$$

The second equation (2) assumes the individual effects to be truly random, i.e., the conditional expectation of the individual effect equals the unconditional expectation of the individual effect – c_i and $\underline{x}_i$ are independent. This assumption is violated if the individual effect is correlated with one or more regressors.

One example for an individual effect in the current context is individual entrepreneurial ability: as van Praag and Cramer (2001) point out, individual entrepreneurial ability is a very important factor for the decision to become an entrepreneur. Since the data does not include a measure of ability, this influence is part of the error term in the model. Furthermore, entrepreneurial ability is likely to be correlated with other variables in the model as a higher ability level is presumably related with more personal income or a higher level of education. However, this means that classical random effects models result in inconsistent estimates because (2) is violated and the c_i are no longer random effects.

A Hausman test can reveal the presence of individual effects which are correlated with regressors in the model. The idea is to compare the coefficients of a random effects model which assumes (2) and a fixed effects model which allows correlation between the individual effect and the regressors. The tested hypothesis is that the difference between the fixed effects coefficients and the random effects coefficients is not systematic. If H_0 is not rejected, (2) holds with a high probability, both coefficient sets are consistent and under further assumptions the random effects estimator is efficient. If H_0 is rejected, most probably only the fixed effects coefficients are consistent. In this study, a Hausman test between linear random effects and fixed effects estimators rejects the H_0 at a <1% significance level ($X^2_{13} = 455.20$). This means that only models which do not assume (2) can estimate coefficients consistently.

There are two prominent ways to deal with the problem of unobserved heterogeneity in panel data models: the most popular way is to cancel out the individual effect (assumed to be constant over time) with fixed effect time demeaning techniques. The general idea is to average every variable in the model over time and then to subtract this average from the original variable value. Consequently, all time-constant effects – including the individual effect – drop out of the equation. Estimation of the model results in unbiased and consistent coefficients for the time-varying effects. The advantage of this model is that no additional assumptions on the distribution of the individual effects have to be made – beside the assumption that the effect is constant over time. The disadvantage is that coefficients for time-constant variables like gen-

der cannot be estimated with a classical fixed effects model. Besides, computation is only possible if the variables vary enough over time. With the given data, only a linear fixed effects regression would be possible; fixed effects Logit estimates do not converge.
However, given that the dependent variable is a binary dummy variable, the model has to be estimated as a model for the probability of someone to become an entrepreneur. Thus, the estimated function should be nonlinear and bound between zero and one – like the normal cumulative distribution function assumed in Probit models.

In this study, I follow another approach to solve the problem of unobserved heterogeneity. This model, also known as *Chamberlain's Random Effects Probit model*, explicitly allows unobservables to be correlated with some of the independent variables under the assumption that the unobserved heterogeneity is normally distributed with a linear mean and a constant variance conditional on the independent variables.[74] In my model, I refer to Mundlak (1978) who restricted the assumption to the means of the independent variables over time:

$$c_i \mid \underline{x}_i \sim N(\psi + \overline{\underline{x}}_i \xi, \sigma_a^2) \tag{3}$$

Thus, the distribution for the individual effects given the explanatory variables is assumed to be normal and depending on the time averages of the time-varying explanatory variables. Unlike in the fixed effects approach, the time-constant variables can still be included in this model. However, in order to distinguish their effect from the individual effect, one has to assume that these time-constant variables are not partially correlated with the unobserved effect. The error term e_{it} is now assumed to be independent Normal(0,1) distributed conditional on ($\underline{x}_i$,a_i), so E($e_i|\underline{x}_i$,a_i)=0. The basic assumption E($u_i|\underline{X}_i$)=0 is no longer violated ($\underline{X}_i$ including all $\underline{x}_i$,a_i) and the estimated coefficients are unbiased and consistent. As the distribution of the $a_i|\underline{x}_i$ is assumed to be Normal(0,σ_a^2) the a_i are classical random effects. Now, the Probit model is (4) and the latent variable behind the Probit model becomes (5).

$$P(y_{it} = 1 \mid x_{it}, c_i) = \Phi(\psi + x_{it}\beta + \overline{x}_i \xi) \tag{4}$$

$$y_{it}^* = \psi + x_{it}\beta + \overline{x}_i \xi + a_i + e_{it} \tag{5}$$

Under the above assumptions, and adding the time averages of the explanatory variables to the equation, I can estimate a classical random effects Probit model by maximum likelihood.[75] The time averages reflect the correlation with the unobserved individual effect while the coef-

[74] See Wooldridge (2002), p. 487 f.
[75] The assumptions have to hold in addition to the original Probit assumptions.

ficients of the x_{it} show the effect of a change in the x_{it} on the probability of becoming self-employed, holding the time average of these x_{it} fixed.

3.3.2 The Data

In this section, the data set and the sample are described, followed by some descriptive statistics on the sample. A comparison of descriptive statistics by employment status concludes.

3.3.2.1 The Mikrozensus Panel

This analysis is based on the German Mikrozensus Panel 2001-2004 which has been published in June 2008.[76] The Mikrozensus is a representative sample of the German population. Participation is enforced by law which leads to a very high participation rate. Once a year the interviewers conduct face-to-face interviews to ask the participants about their social and economic situation, their education and the job market.

Until 2006 the Mikrozensus data has only been published as cross-sectional data. With the Mikrozensus Panel 1996-1999 and the Mikrozensus Panel 2001-2004 the panel structure of the data can be used. Every year one quarter of the panel is being replaced which leads to a four-year rotating panel with about 195,520 participants with 500,181 observations in total for the Mikrozensus Panel 2001-2004.

Until 2004, the reference period for the largest part of the questions was a week, mostly in the end of April. After that, survey design switched to continuous interviewing during the year and an "average" week as reference.[77] As the Mikrozensus Panel 2001-2004 refers to the system of reporting one week in April, there is no evidence in the data for persons who switch to another employment status and back between two interviews. Short-term self-employment as well as short-term unemployment between two interviews is not reflected in the data.

3.3.2.2 The Sample

The sample is restricted to individuals who are between 16 and 65 during the whole time of the panel, i.e. not younger than 16 in 2001 because those are very likely to still be in education and not older than 65 in 2004 because 65 was the official age of retirement in 2001-2004 (331,828 observations dropped). I exclude farmers and fishermen (3,917 observations), civil servants (37,735 obs.) and individuals who are currently in education (35,161 obs.), vocational training, and military service (1,941 obs.). As these individuals base their occupational decisions on different determinants or only have a limited set of occupational choice they

[76] This preliminary version of the Mikrozensus Panel still lacks longitudinal weights. The full version will be published in autumn 2008.

[77] See Hansch (2006) for more details on the subject.

would distort the analysis. Helping family members (1,953 obs.) are also excluded from the analysis as they are not self-employed in a sense that they run their own business.[78] This leaves a total of 234,596 observations in four years.

3.3.2.3 Descriptive Statistics of the Sample

While table 3.1 provides a description of the variables in the sample, table 3.2 gives a summarizing overview of the characteristics of the individuals in the sample. For simplification, only the 2003 sample is shown.[79]

2003	Description
Socio Variables	
female	Dummy for females
age	Age in years (year-birth year)
single	Dummy for single individuals
married	Dummy for married individuals
widowed	Dummy for widowed individuals
divorced	Dummy for divorced individuals
famstatus	Family status: 1 if single, 2 if married, 3 if widowed, 4 if divorced
kids	Dummy for having kids who live in the family
hhcount	Number of persons in the household
east	Dummy for individuals who live in East Germany
Education	
noeduc	Dummy for individuals who do not have any school degree
higheduc	Dummy for individuals who finished university or an applied sciences university or have a higher degree
educ	Dummy for individuals who graduated from a secondary school
Financial Variables	
persincome	Personal net income in 1,000 Euro per month
highincome	Dummy for individuals whose personal net income is above the median in the population
hhincome	Household net income in 1,000 Euro per month
hhpercap	Household net income per capita (income divided by the square root of household members)
home	Dummy for individuals who have income from real estate or who lived in their own apartment / house in 2002
incomeassets	Dummy for individuals who have income from liquid assets
Job Variables	
employed	Dummy for individuals who are employed according to ILO standards
UE	Dummy for individuals who are unemployed according to ILO standards
inactive	Dummy for individuals who are inactive according to ILO standards
SE	Dummy for self-employed individuals
SE1	Dummy for individuals whose primary occupation is self-employment
SE2	Dummy for individuals whose secondary occupation is self-employment
NewSE	Dummy for individuals who are not self-employed in the current year but will be next year
dsearch	Duration of job search in months
dsearch_high	Dummy for individuals who have been looking for a job for longer than 1 year

Table 3.1: Variable Descriptions

In 2003, 49.5% of the individuals are females, 17.1% come from East Germany. The average age is at 44 years. 25.1% are single, 64.3% married, 7.9% divorced and 2.7% widowed. 66% have at least one child who lives in the family. The average number of persons in a household is 2.7. 2.7% of the sample have not graduated from any school, 82.4% possess some qualify-

[78] For more information on the restrictions to the data set see Caliendo et al (2006).
[79] The 2003 data is the most recent complete observation set for all variables including NewSE.

ing degree, and 11.1% have at least completed a university degree or a degree at a university of applied sciences.

2003	Obs	Mean	Std. Dev.	Min	Max
Socio Variables					
sex	59,514	0.495	.500	0	1
age	59,514	44.018	12.102	18	64
single	59,514	0.251	.434	0	1
married	59,514	0.643	.479	0	1
divorced	59,514	0.079	.270	0	1
widowed	59,514	0.027	.161	0	1
famstatus	59,514	1.934	.768	1	4
kids	59,514	0.662	.473	0	1
hhcount	59,298	2.689	1.237	1	9
east	59,514	0.171	.376	0	1
Education					
noeduc	56,797	0.028	0.164	0	1
higheduc	55,151	0.111	0.314	0	1
educ	56,086	0.824	0.381	0	1
Financial Variables					
persincome	56,966	1,265.863	1,324.559	0	18,000
highincome	56,966	0.584	0.493	0	1
hhincome	56,808	2,535.814	1,813.665	75	18,000
hhpercap	56,808	1,590.511	1,103.144	33.541	18,000
home	59,514	.410	.492	0	1
incomeassets	59,514	.029	.168	0	1
Job Variables					
employed	59,514	0.650	0.477	0	1
UE	59,514	0.111	0.314	0	1
inactive	59,514	0.239	0.426	0	1
SE	59,514	0.088	0.283	0	1
SE1	59,514	0.082	0.274	0	1
SE2	59,514	0.008	0.088	0	1
NewSE	51,132	0.013	0.115	0	1
dsearch	7,278	19.286	16.913	1	48
dsearch_high	7,278	0.501	0.500	0	1

Table 3.2: Summary Statistics 2003

In the Mikrozensus, financial variables are mainly obtained from the question

> *"What is your / your household's monthly net income? Please count all forms of income, i.e. salary, income from self-employment, pensions, public subsidies, income from renting and leasing, housing and family allowances"*.

The average personal net income in 2003 is 1,266 Euro per month; the average household net income is at 2,536 Euro. Mean household net income per capita – defined as the total net household income adjusted by a standard equivalence scale[80] – is at 1,591 Euro per month.

As the Mikrozensus does not provide an explicit variable for assets I construct a dummy variable for individuals who possess assets which they could use as collateral if they wanted to borrow money: this variable equals one if an individual states some income from renting and leasing in the Mikrozensus. Besides, individuals who live in their own house/apartment are defined to have assets. Unfortunately, the question whether a person owns the house or apartment she lives in has only been asked in the 2002 wave. Therefore, I assign anyone who, in 2002, owns the apartment they live in to have assets in the 2001 and 2003 years as well. In

[80] Income divided by the square root of the number of household members.

2003, 41.0% of the sample own a house or apartment. In addition to that, wealth in the form of liquid assets is measured with a dummy variable for those who state to have some income from financial assets. In 2003, this was 2.9% of the population.

The employment status (employed, unemployed, and inactive) follows the ILO standard. In 2003, 65% of the persons in the sample are employed, 11% are unemployed, and 24% are part of the inactive population.

An individual is classified as being self-employed when she reports self-employment as her primary or secondary activity.[81] In 2003, 8.8% of the sample are self-employed, 8.2% as primary occupation, 0.8% as a secondary occupation (some individuals report self-employment in both primary and secondary occupation). A "New Self-Employed" (NewSE) is a person who is not self-employed in the current year and is self-employed in the following year. 680 such transitions can be found in the panel period, 1.3% of the 2003 sample is NewSE.

The variable "dsearch" counts the months an individual has already been looking for a job. For the 7,278 unemployed individuals in the 2003 sample the mean duration of job search is 19.3 months.

3.3.2.4 Characteristics by Employment State

An analysis by employment state comes to important differences between self-employed, employed, unemployed and inactive persons as shown in table 3.3.

Only 29.4% of the self-employed are female compared to 44.9% of the employed, 44.4% of the unemployed, and 69.8% of the inactive population. As much as 28.7% of the self-employed are very well educated (at least a degree at a university of applied sciences), compared to 11.2% of the employed, 6.6% of the unemployed, and 6.4% of the inactive population. Self-employed individuals are much less likely to have no education at all than the other groups. 92.8% have some qualifying degree.

Employed and unemployed persons are significantly younger than the self-employed in the sample whereas, on average, the inactive population is older.

With an average of 2,482 Euro per month, personal income is substantially higher for the self-employed than for the employed (1,485€), unemployed (605€) and the inactive population (628€). The same is true for the household income. However, while the personal income does not differ between inactive and unemployed individuals, inactive persons have a much higher

[81] Later in the analysis, I will distinguish between being self-employed as a primary and as a secondary activity.

household income than unemployed individuals, supposedly because their spouses and family are working.

2003	Self-Employed	Employed	Unemployed	Inactive
Socio Variables				
female	0.294	0.449***	0.444***	0.698***
age	45.493	40.972***	41.738***	51.695***
single	0.215	0.297***	0.352***	0.109***
married	0.679	0.615***	0.495***	0.764***
widowed	0.015	0.015	0.021**	0.061***
divorced	0.091	0.073***	0.132***	0.065***
kids	0.683	0.705***	0.713***	0.529***
hhcount	2.730	2.738	2.543***	2.626***
east	0.139	0.159***	0.340***	0.133
Education				
noeduc	0.006	0.018***	0.044***	0.051***
higheduc	0.287	0.112***	0.066***	0.064***
educ	0.928	0.869***	0.746***	0.710***
Financial Variables				
persincome	2,482.266	1,485.056***	604.679***	628.116***
highincome	0.823	0.752***	0.179***	0.293***
hhincome	3,715.239	2,695.353***	1,516.068***	2,205.439***
hhpercap	2,336.808	1,685.408***	955.242***	1391.263***
home	0.560	0.399***	0.224***	0.468***
incomeassets	0.061	0.024***	0.019***	0.033***

Table 3.3: Mean Characteristics by Employment State
Stars indicate whether the mean is significantly different from the mean in the self-employed sample (two-sample t-test with equal variances, ***/**/* for significance at 1%/5%/10% level).

Finally, the share of self-employed who own a house or apartment is significantly larger than the shares in the other groups (56.0% vs. 40.0%, 22.4% and 46.8%, respectively). The same is true for liquid assets (6.1% vs. 2.4%, 1.9% and 3.3%, respectively).

In addition to the variables described in table 3.1, I will use a few interactions between variables in my estimations, such as the common influence of unemployment and gender, unemployment and being from East Germany or interactions of variables and year dummies.

3.4 Estimation Results

In this section, I present estimation results for the whole sample, the sample split by gender, and for the sub sample of the unemployed. I compare the results to a classical random effects model and to the results of previous studies on the subject.

3.4.1 Whole Sample

The basic model, summarized in table 3.4, contains measures for gender, age, squared age, East Germany, education, marital status, income and assets, the current job situation and

dummies for the years 2002 and 2003, 2001 being the reference year. The dependent variable, NewSE, refers to both self-employed as a primary and as a secondary occupation.

Women are significantly less likely to become self-employed than men – as are individuals from East Germany compared to those from West Germany. The likelihood to become self-employed rises significantly with age, but at a decreasing rate. In two regressions with age dummies, one with a 16 to 25, a 26 to 40 and a 40 to 65 group, the other with 16 to 29, 30 to 49 and 50 to 65 years, no dummy can be shown to be of significant influence. The same is true for regressions with variables for the number of persons in the household and a dummy for kids in the household.[82]

Married persons are more likely to switch to self-employment than singles, but the coefficients for widowed and divorced individuals are not significant. A high level of education significantly decreases the probability of becoming self-employed – presumably because a university degree facilitates the entry to the regular wage market and thus deters the well educated from self-employment.

Personal assets in the form of house ownership or income from liquid assets do not affect the decision to become self-employed in any way. This would reject Knight's view that assets are important in order to build an enterprise. The large amount of governmental aid for new entrepreneurs, however, might in fact act as a subsidy of accumulated wealth.

While the coefficient for personal net income is significantly negative, the net household income per capita has a positive effect on the decision. One reason might be that a higher income of the spouse influences the decision in favour of becoming self-employed as the spouse bears part of the financial risk for the family.

As the unemployed and inactive population is far more likely to become self-employed than the employed population the push hypothesis – which claims that unemployment pushes people into self-employment – is affirmed.

Between 2003 and 2004, significantly fewer persons decide to become self-employed than between 2001 and 2002 – despite the new legislation in 2003 which establishes more subsidies for the unemployed who found new firms.

However, if interactions between job status and the years are included, it becomes obvious that the unemployed and inactive population is significantly more likely to switch to self-employment between 2003 and 2004 than in the other periods. The new legislation seems to

[82] Regressions with age dummies, hhcount and kids are not shown in the table.

have affected the unemployed and inactive, while the employed became more hesitant in this period.

	original	UE_year, inactive_year	UE_east, inactive_east	UE_sex, inactive_sex
Social Variables				
female	-0.150***	-0.151***	-0.149***	-0.138***
	(0.0274)	(0.0274)	(0.0274)	(0.0316)
age	0.367***	0.377***	0.368***	0.369***
	(0.0729)	(0.0738)	(0.0729)	(0.0729)
age_sq	-0.00326***	-0.00358***	-0.00327***	-0.00330***
	(0.000762)	(0.000780)	(0.000762)	(0.000763)
east	-0.127***	-0.126***	-0.0439	-0.127***
	(0.0336)	(0.0337)	(0.0399)	(0.0337)
higheduc	-0.256***	-0.256***	-0.254***	-0.256***
	(0.0882)	(0.0884)	(0.0883)	(0.0883)
dfamstatm	0.291**	0.292**	0.294**	0.291**
	(0.136)	(0.136)	(0.136)	(0.136)
dfamstatw	0.0466	0.0379	0.0421	0.0448
	(0.308)	(0.309)	(0.308)	(0.308)
dfamstatd	0.0121	0.0141	0.0115	0.0109
	(0.150)	(0.150)	(0.150)	(0.150)
p-value	0.0000	0.0000	0.0000	0.0000
Financial Variables				
p_income	-0.00964***	-0.00942***	-0.00958***	-0.00952***
	(0.00243)	(0.00243)	(0.00243)	(0.00243)
h_percap	0.00627**	0.00605**	0.00631**	0.00623**
	(0.00272)	(0.00272)	(0.00272)	(0.00272)
home	0.101	0.102	0.0976	0.100
	(0.173)	(0.173)	(0.173)	(0.173)
incomeassets	-0.0716	-0.0719	-0.0744	-0.0759
	(0.0847)	(0.0848)	(0.0848)	(0.0848)
p-value	0.0016	0.0022	0.0018	0.0019
Employment Variables				
UE	1.616***	1.506***	1.676***	1.636***
	(0.0629)	(0.0838)	(0.0724)	(0.0815)
UE_year02		0.0297		
		(0.108)		
UE_year03		0.317***		
		(0.108)		
UE_east			-0.249*	
			(0.141)	
UE_female				-0.0345
				(0.124)
inactive	1.690***	1.584***	1.685***	1.842***
	(0.0714)	(0.0874)	(0.0748)	(0.128)
inactive_year02		0.130		
		(0.0950)		
inactive_year03		0.224**		
		(0.103)		
inactive_east			0.150	
			(0.241)	
inactive_female				-0.219
				(0.152)
p-value	0.0000	0.0000	0.0000	0.0000
Year Dummies				
dy02	-0.0556	-0.0588	-0.0581	-0.0557
	(0.0378)	(0.0411)	(0.0379)	(0.0378)
dy03	-0.161***	-0.181***	-0.165***	-0.162***
	(0.0597)	(0.0627)	(0.0597)	(0.0597)
p-value	0.0151	0.0086	0.0134	0.0148
Mean Values				
…	…	…	…	…
p-value	0.0000	0.0000	0.0000	0.0000
Constant	-2.885***	-2.861***	-2.903***	-2.886***
	(0.193)	(0.193)	(0.193)	(0.194)
Observations	120,017	120,017	120,017	120,017
Number of fid	56,444	56,444	56,444	56,444

Table 3.4: Regression Results: Whole Sample
(standard errors in brackets, mean values included but not shown in the table, ***/**/* for significance at 1%/5%/10% level).

Interactions between job status and East Germany reveal that being unemployed in East Germany lowers the likelihood of becoming self-employed in addition to the negative effect of being from East Germany – an effect which is not significant for the inactive population.
Common effects of gender and the job situation, reflected by interactions between gender and job status, are not significant – nor are interactions between gender and marital status, or between East Germany and home ownership. In a regression on all employed persons, the current branch of employment, reflected by 6 dummy variables, cannot be shown to have any influence on the decision to become self-employed.
F-tests for the included mean values[83] reject the hypothesis that the coefficients of the values are jointly zero with p-values below 0.0001. This implies that the influence of the unobserved heterogeneity, represented via the mean values, is significantly different from zero.

Results are similar if only the self-employed as a primary occupation are concerned (results not shown in the table). Unfortunately, the number of observations for the self-employed as a secondary occupation is not large enough to estimate a separate model.[84]
A separate estimation for East Germany is not possible either as the number of observations for new entrepreneurs lies between 71 and 107.[85] Separate results for West Germany equal the results for the whole sample (results not shown in the table).

	original	UE_year, inactive_year
female	-0.0038***	-0.0037***
age (+1 year)	0.0028***	0.0049***
east	-0.0031***	-0.0029***
higheduc	-0.0060***	-0.0058***
home	0.0027	0.0026
incomeassets	-0.0018	-0.0017
UE	0.1667***	0.1411***
inactive	0.1816***	0.1564***
UE_year03		0.0109***
inactive_year03		0.0071***
p_income (+100€)	-0.0003***	-0.0002***
p_income (+1%)	-0.0001***	-0.00005***
p_income (+100%)	-0.0037***	-0.0036***
h_percap (+100€)	0.0002**	0.0002**
h_percap (+1%)	0.00004**	0.00003**
h_percap (+100%)	0.0042**	0.0039**

Table 3.5: Marginal Effects: Whole Sample
Effect of a discrete change of a variable from 0 to 1 – if not indicated otherwise.
Effects averaged across the sample. Effect of age includes effect of age_sq.
Significances according to estimation in Table 3.4.

[83] See chapter 3.3 for the model description.
[84] The number of observations is 191, 231 and 193 in the years 2001, 2002 and 2003.
[85] The number of observations is 71, 87 and 107 in the years 2001, 2002 and 2003.

Table 3.5 shows marginal effects for some of the variables for the original model and the model with an interaction between job status and the year. The marginal effects of a change in the regressors on the probability of becoming self-employed are averaged across the population. In the original model, being a woman or being from East Germany decreases the probability of becoming self-employed ceteris paribus by 0.4 and 0.3 percentage points. As far as age is concerned, an additional year leads to an increase of 0.3 percentage points. If personal net income is increased by 100€ per month, the probability of becoming self-employed is decreased by 0.03 percentage points while the same raise in household net income per capita increases the probability by 0.02 percentage points. The largest marginal effects arise when job status is concerned: ceteris paribus, being unemployed increases the switching probability by 16.7 percentage points while being inactive has an effect of 18.16 percentage points.
In the model which includes interactions between job status and the year the marginal effects of the interaction terms can be calculated: the average effect of being unemployed in 2003 is at 1.1 percentage points while the effect of being inactive in 2003 amounts to 0.7 percentage points.

3.4.2 Results by Gender

Separate estimations for the two genders – as shown in table 3.6 – show that while the influence of age, squared age, and East Germany on the decision to become self-employed is similar for both sexes, the effect of the financial situation on the decision differs between men and women. Neither personal nor household income per capita nor home ownership or having income from liquid assets is of significant influence to a male's decision to become self-employed whereas personal net income affects the women's decision in a negative and household income per capita affects it in a positive way. Another difference is revealed in the effect of the level of education: only well educated men are less likely to found their own business as opposed to the women for whom the coefficient of high education is negative but insignificantly so.
The year dummies for 2002 and 2003 are significantly negative for men but not significant for women. As interactions with the current job situation are concerned, results differ between sexes in the interaction of unemployment and the year dummies: while male unemployed persons are significantly more likely to become self-employed between 2003 and 2004, this is not the fact for the female, unemployed share of the population.

	Male original model	Female original model	Male UE_year, inactive_year	Female UE_year, inactive_year	Male UE_east, inactive_east	Female UE_east, inactive_east
Social Variables						
age	0.438***	0.271**	0.443***	0.275**	0.439***	0.278**
	(0.0940)	(0.114)	(0.0966)	(0.115)	(0.0941)	(0.115)
age_sq	-0.00366***	-0.00272**	-0.00387***	-0.00303**	-0.00367***	-0.00279**
	(0.000981)	(0.00119)	(0.00102)	(0.00122)	(0.000982)	(0.00120)
east	-0.132***	-0.126**	-0.132***	-0.123**	-0.0346	-0.0781
	(0.0439)	(0.0519)	(0.0439)	(0.0521)	(0.0510)	(0.0632)
higheduc	-0.392***	-0.0396	-0.392***	-0.0403	-0.391***	-0.0380
	(0.113)	(0.136)	(0.113)	(0.137)	(0.113)	(0.137)
dfamstatm	0.241	0.348	0.237	0.359	0.243	0.354
	(0.172)	(0.220)	(0.172)	(0.221)	(0.172)	(0.221)
dfamstatw	0.0306	0.106	0.0230	0.107	0.0134	0.108
	(0.508)	(0.388)	(0.507)	(0.390)	(0.506)	(0.390)
dfamstatd	0.0317	0.00313	0.0255	0.0157	0.0330	0.00355
	(0.197)	(0.232)	(0.197)	(0.233)	(0.197)	(0.234)
p-value	0.0000	0.0232	0.0000	0.0201	0.0000	0.0864
Financial Variables						
p_income	-0.00491	-0.0222***	-0.00459	-0.0220***	-0.00478	-0.0222***
	(0.00324)	(0.00474)	(0.00325)	(0.00476)	(0.00324)	(0.00476)
h_percap	0.00314	0.00890**	0.00277	0.00877**	0.00312	0.00890**
	(0.00420)	(0.00347)	(0.00421)	(0.00348)	(0.00420)	(0.00348)
home	0.0555	0.226	0.0561	0.226	0.0523	0.214
	(0.218)	(0.284)	(0.218)	(0.284)	(0.218)	(0.285)
incomeassets	-0.131	0.0492	-0.132	0.0521	-0.138	0.0469
	(0.105)	(0.140)	(0.106)	(0.141)	(0.106)	(0.141)
p-value	0.3151	0.0002	0.3445	0.0002	0.3181	0.0002
Employment Variables						
UE	1.660***	1.516***	1.530***	1.435***	1.734***	1.572***
	(0.0823)	(0.108)	(0.109)	(0.137)	(0.0952)	(0.122)
UE_year02			0.0427	0.0284		
			(0.136)	(0.174)		
UE_year03			0.385***	0.213		
			(0.139)	(0.170)		
UE_east					-0.298	-0.213
					(0.185)	(0.216)
inactive	1.878***	1.543***	1.758***	1.461***	1.949***	1.517***
	(0.130)	(0.0990)	(0.170)	(0.113)	(0.139)	(0.101)
inactive_year02			0.164	0.122		
			(0.172)	(0.117)		
inactive_year03			0.166	0.198		
			(0.188)	(0.126)		
inactive_east					-0.487	0.486
					(0.389)	(0.304)
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Year Dummies						
dy02	-0.0855*	-0.0141	-0.0859	-0.0243	-0.0873*	-0.0176
	(0.0493)	(0.0578)	(0.0524)	(0.0658)	(0.0493)	(0.0581)
dy03	-0.265***	-0.0204	-0.287***	-0.0272	-0.267***	-0.0248
	(0.0787)	(0.0897)	(0.0822)	(0.0953)	(0.0788)	(0.0902)
p-value	0.0012	0.9689	0.0007	0.9337	0.0011	0.9533
Mean Values						
…	…	…	…	…	…	…
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Constant	-2.495***	-3.495***	-2.470***	-3.480***	-2.515***	-3.521***
	(0.249)	(0.346)	(0.249)	(0.347)	(0.249)	(0.348)
Observations	58274	61743	58274	61743	58274	61743
Number of fid	27420	29024	27420	29024	27420	29024

Table 3.6: Regression Results: Results by Gender
(standard errors in brackets, mean values included but not shown in the table,
***/**/* for significance at 1%/5%/10% level).

3.4.3 Duration of Unemployment

In order to determine the effects of the duration of unemployment on the likelihood to become self-employed, I estimate a separate model on the subsample of the unemployed population. The reason is that the duration of unemployment is recorded only for the unemployed – with a missing value for the rest of the population. Results are shown in table 3.7.

As in the whole sample, women and East Germans are less likely to become self-employed than men and West Germans, respectively. Age, however, is no longer of significant influence on the decision once the sample is restricted to the unemployed population. While high education no longer plays a role, the marital status seems to be very important for the decision process: married, widowed and divorced persons are more likely to become self-employed than singles.

Again, personal net income is negatively correlated with the probability of becoming self-employed. However, the coefficient of household net income per capita is no longer significant in the sub sample of the unemployed. It is interesting to see that while home ownership in general still does not influence the decision significantly, it does influence unemployed persons from East Germany in a negative way. Furthermore, income from liquid assets is a positive contributor to the likelihood of becoming self-employed but only until unemployment benefits are added to the regression.

Even though subsidies for new firm foundation are often linked to the unemployment benefits the new entrepreneur received before becoming self-employed, unemployment benefits do not play a significant role in the decision process.

If an unemployed individual has been looking for a job for longer than a year (the dummy dsearch_high equals one), she is much more likely to create her own business than with a shorter record of unemployment. This might be the reason for the high number of firm foundations out of dire straits in Germany.[86]

As compared to 2001, the influence of a long job search on the decision to become self-employed is significantly lower in 2003. Here, the introduction of the EGZ is a potential explanation because this incentive encouraged the unemployed to create a new business, regardless of how long they had already been looking for a job.

[86] 63% of the former unemployed state they created their business out of dire straits: see KfW (2008) for more information.

	Original Model	home_east	dsearch_h*year	with UE_help
Social Variables				
female	-0.149**	-0.143**	-0.149**	-0.122*
	(0.0678)	(0.068)	(0.0680)	(0.0715)
age	-0.1000	-0.0813	-0.130	-0.0720
	(0.194)	(0.19)	(0.195)	(0.204)
age_sq	0.000745	0.000595	0.00125	0.000270
	(0.00213)	(0.0021)	(0.00216)	(0.00226)
east	-0.294***	-0.331***	-0.290***	-0.275***
	(0.0720)	(0.093)	(0.0722)	(0.0756)
higheduc	-0.290	-0.287	-0.282	-0.275
	(0.241)	(0.24)	(0.241)	(0.255)
dfamstatm	0.780**	0.779**	0.757**	0.857**
	(0.346)	(0.35)	(0.348)	(0.368)
dfamstatw	1.625**	1.623**	1.525**	2.081**
	(0.771)	(0.77)	(0.776)	(0.836)
dfamstatd	0.767**	0.772**	0.770**	0.816**
	(0.355)	(0.36)	(0.356)	(0.376)
p-value	0.0002	0.0010	0.0002	0.0008
Financial Variables				
p_income	-0.0433***	-0.0443***	-0.0429***	-0.0546***
	(0.0117)	(0.012)	(0.0117)	(0.0134)
h_percap	0.00436	0.00468	0.00400	0.0178
	(0.0112)	(0.011)	(0.0112)	(0.0127)
home	0.365	0.672	0.363	0.207
	(0.567)	(0.59)	(0.570)	(0.647)
home_east		-3.497**		
		(1.68)		
incomeassets	0.421*	0.536**	0.445*	0.00588
	(0.253)	(0.27)	(0.254)	(0.293)
incomeassets_east		-1.093		
		(0.82)		
UEhelp				0.180
				(0.213)
p-value	0.0004	0.0002	0.0004	0.0030
Employment Variables				
dsearch_high	0.359***	0.358***	0.507***	0.369***
	(0.116)	(0.12)	(0.169)	(0.121)
dsearch_high_02			-0.213	
			(0.196)	
dsearch_high_03			-0.329*	
			(0.190)	
p-value	0.0020	0.0020	0.0196	0.0020
year dummies				
dy02	0.0207	0.0171	0.000354	0.0316
	(0.101)	(0.10)	(0.121)	(0.108)
dy03	0.155	0.147	0.227	0.170
	(0.142)	(0.14)	(0.151)	(0.151)
p-value	0.3532	0.3808	0.1128	0.3664
Mean values				
…				
p-value	0.0000	0.0000	0.0000	0.0000
Constant	-4.389***	-4.366***	-4.453***	-4.444***
	(0.503)	(0.50)	(0.511)	(0.535)
Observations	11,854	11,854	11,854	11,242
Number of fid	8,360	8,360	8,360	7,903

Table 3.7: Regression Results: Unemployed
(standard errors in brackets, mean values included but not shown in the table,
***/**/* for significance at 1%/5%/10% level).

3.4.4 Alternative Model Specification

In this section, I will compare the results of the model I chose to a classical random effects model without included mean values.

In addition to the assumption of strict exogeneity (1), classical random effects models assume the individual effects to be truly random (2), i.e., the conditional expectation of the individual effect equals the unconditional expectation of the individual effect – c_i and $\underline{x}_i$ are independent. However, if (2) is violated, the estimated coefficients are not consistent. The model comparison can be seen in table 3.8:

	Chamberlain's RE Probit	Traditional RE Probit
Social Variables		
female	-0.150***	-0.00884***
	(0.0274)	(0.00124)
age	0.367***	0.00316***
	(0.0729)	(0.000361)
age_sq	-0.00326***	-4.04e-05***
	(0.000762)	(4.12e-06)
east	-0.127***	-0.00724***
	(0.0336)	(0.00148)
higheduc	-0.256***	0.0143***
	(0.0882)	(0.00163)
dfamstatm	0.291**	0.00179
	(0.136)	(0.00146)
dfamstatw	0.0466	-0.000336
	(0.308)	(0.00347)
dfamstatd	0.0121	0.000418
	(0.150)	(0.00219)
p-value	0.0000	0.0000
Financial Variables		
p_income	-0.00964***	0.000273***
	(0.00243)	(5.91e-05)
h_percap	0.00627**	0.000254***
	(0.00272)	(5.94e-05)
home	0.101	0.00799***
	(0.173)	(0.00118)
incomeassets	-0.0716	0.00387*
	(0.0847)	(0.00224)
p-value	0.0016	0.0000
Employment Variables		
UE	1.616***	0.0181***
	(0.0629)	(0.00132)
inactive	1.690***	0.00347***
	(0.0714)	(0.00126)
p-value	0.0000	0.0000
Year Dummies		
dy02	-0.0556	0.00544***
	(0.0378)	(0.000645)
dy03	-0.161***	0.00845***
	(0.0597)	(0.000664)
p-value	0.0151	0.8552
Mean values		
…	…	-
p-value	0.0000	-
Constant	-2.885***	-0.0491***
	(0.193)	(0.00728)
Observations	120,017	120,017
Number of fid	56,444	56,444

Table 3.8: Model Comparison
(standard errors in brackets, mean values included but not shown in the table, ***/**/* for significance at 1%/5%/10% level).

Results are quite similar for both models as far as the influence of gender, East Germany, and the current job situation is concerned.

The effect of a higher education changes from significantly negative in my model to significantly positive in the random effects model. The variable for high education obviously takes up part of the correlation with the individual effects in the random effects model.

The most interesting differences can be found in the set of financial variables: while in my model, personal net income significantly affects the decision to become self-employed in a negative way, the coefficient is not significant in the random effects model. Home ownership and income from assets, however, are of significant positive influence in the random effects model only. The effect of household net income per capita is the same for both models.

To sum things up: the question whether assumption (2) holds is crucial to the results, especially for the influence of high education and the financial situation on the decision to become self-employed.

3.4.5 Comparison with previous Studies

Generally, my results are consistent with the findings of Glocker and Steiner (2007) and Caliendo, Fossen and Kritikos (2006). For Germany, they also find a positive effect of unemployment on the decision to become self-employed and higher self-employment rates for men than for women.

Concerning assets, my study is in line with Holtz-Eakin, Joulfaian and Rosen (1994), who analyze the impact of inheritances, liquid assets and home ownership on the transitions into self-employment. They find insignificant influence of home ownership and liquid assets on the probability of transition. Evans and Leighton (1989), who find a strong positive influence of assets on the transition decision, do not differentiate between the sources of wealth. However, they also find that personal income affects the decision in a negative way – as do Evans and Jovanovic (1989), and Meyer (1990).

My results concerning personal unemployment as a positive contributor to the decision to become self-employed are consistent with a large number of studies, e.g. Evans and Leighton (1989) for the U.S. or Constant and Zimmermann (2004) for Germany.

3.5 Discussion

This study's main contribution is twofold: first of all, the factors which influence the transition into self-employment are analyzed based on large panel data set. This is the first study to use the Mikrozensus Panel 2001-2004. Panel data offers the possibility to analyze individual

transitions instead of comparing consecutive cross sectional stock numbers. Besides, the advantage of panel data is that the influence of unobserved individual effects can be included in an analysis.

Cross sectional studies have to focus on the probability of an individual to be self-employed at one point in time. The conditions that determine the actual transition into self-employment cannot be analyzed in a cross sectional setting. The most puzzling example for endogeneity issues in this context is the influence of assets on the decision to become self-employed: are persons with assets more likely to become self-employed – or does self-employment simply put a person in the position to accumulate more wealth? This question cannot be answered based on cross sectional data where either the already self-employed OR those who wish to become self-employed can be observed. In this panel data set, the individuals are observed before becoming self-employed AND after the transition which enables me to analyze wealth before an actual transition and, more generally speaking, the dynamics behind this transition. Concerning the EGZ it is obvious that the pure effect of its introduction can not be assessed in this study because of the lack of a reference group. However, I am able to identify a strong positive effect of being unemployed in 2003 on the decision to become self-employed, holding the other personal factors constant.

The Mikrozensus Panel 2001-2004 is a very useful data set in terms of the number of observations. However, the four year time horizon is too short. First of all, the influence of complete business cycles on the individuals' behaviour cannot be analyzed with this data set. This is why national unemployment rates or GDP are not included in my models. Besides, this short period of time makes it impossible to monitor the new self-employed persons in a long-term study. When the individual is not self-employed in the first period and becomes self-employed in the second period, a maximum of two periods is left to monitor this person's success – assuming the person enters self-employment in 2002 and stays in the sample until 2004. This is the case for 444 observations in the sample. In order to assess the long-term effects of the EGZ, a panel data set with at least four years after its introduction in 2003 would be necessary because the subsidy can be paid for up to three years and the real success of the receivers can only be identified after this period.

Furthermore, the data set could be improved by including some measure of risk aversion as most individual decisions are affected by risk attitudes. Here, a time constant individual risk aversion is assumed to be part of the unobserved individual effect. However, an explicit

analysis of the correlation between risk aversion and the transition into self-employment as in Wagner (2002) is not possible in this study.
Finally, the Mikrozensus lacks usable asset variables. While I included a dummy variable for home ownership and a dummy for those who have some income from assets – both of which turn out not to be of significant influence on the decision to become self-employed – some specified level of wealth would be a large amelioration of the data.

As far as the model I use is concerned, the advantage of my approach is the consideration of the unobserved individual effects on the one side and the nonlinearity on the other side. The assumption of a conditional normal distribution of the individual effects might seem strong as compared to no distributional assumption in a fixed effects model. However, with the given data, the latter would only have been feasible in its linear version. This is why I preferred Chamberlain's Random Probit model.

3.6 Policy Implications

The analysis shows that the unemployed are more likely to found their own business than the employed. As far as governmental subsidies are concerned, the analysis shows that they seem to be successful. In 2003, the year when an additional subsidy for the unemployed was introduced, the likelihood for an unemployed to become self-employed rose significantly according to my results. 95,000 unemployed received the new subsidy during the first year after its introduction.[87] However, the analysis can only show that unemployed persons in 2003 were more likely to switch to self-employment than the unemployed persons in 2001 and 2002. Whether the reason is to be seen in the new EGZ or an underlying general improvement of the economic conditions in Germany can not be identified with this analysis. The reason is that all unemployed were equally affected by the introduction of the EGZ, i.e., there is no real reference group of unemployed Germans who were not entitled to the new subsidy.
While we know that the share of new self-employed is higher among the unemployed than among the employed, the question of whether this discrepancy really improves the economic development remains unanswered. The 2003 subsidy was not bound to any proof of economic viability or feasibility – one of the reasons the government restricted the access to the EGZ in 2004 and 2005 and introduced a new law in 2006. While Caliendo, Steiner and Baumgartner (2007) are quite enthusiastic about the success of the previously unemployed new entrepre-

[87] See KfW (2008) for more details.

neurs, KfW (2008) points to the fact that in 2007, only one third of all new firm foundations by previously unemployed persons were motivated by a new idea, two thirds were foundations out of dire straits. They also find that foundations by unemployed persons have a lower chance of survival: after two or three years, 41% have stopped their business as compared to 31% of all new entrepreneurs.

According to KfW (2008), business foundation by employees is far more successful than foundation by unemployed. But labour earnings, which can also be interpreted as the opportunity cost of becoming self-employed, can act as a deterrent to self-employment entry. An employee has more to loose by becoming self-employed. My study confirms this view by significantly negative coefficients of personal net income throughout all regressions. A similar argument can explain the negative influence of a high level of education on the probability of transition: as a high level of education facilitates the entry into the wage market and leads to higher wages, this again increases the opportunity cost of self-employment. The according policy implication would be to create more incentives to employees to become self-employed and to expand activities at universities which inform about the chances of self-employment.

3.7 Conclusion

This paper analyzes the characteristics of the individuals who become self-employed. Based on 700 transitions into self-employment per year in the Mikrozensus Panel 2001-2004, I estimate a Chamberlain's Random Effects Probit model explicitly accounting for unobserved heterogeneity by including individual means as controls for the unobserved individual effects.
Besides the finding that male, unemployed or inactive persons are more likely to become self-employed than the rest of the population, personal wealth cannot be shown to affect the decision significantly. While household income has a positive effect on the probability of becoming self-employed, a high personal income decreases it. Finally, a high level of educations is shown to negatively affect the decision to switch into self-employment. It can be shown that accounting for unobserved individual effects leads to opposite estimates of education and financial situation than a classical random effects model.
Being unemployed in 2003 – and thus presumably the introduction of the EGZ in 2003 – has a positive effect on the probability of becoming self-employed, holding the other influencing variables constant. While this means that the government seems to succeed in encouraging the unemployed and inactive to become self-employed, it should keep an eye on the employed,

well paid and highly educated share of the population. Whereas the latter is clearly underrepresented in the group of new entrepreneurs, it could be shown that these new entrepreneurs are more successful and contribute more to the economic development than the previously unemployed and inactive.

Even though the Mikrozensus Panel 2001-2004 offers a large number of observations, the four year time horizon is too short for analyzing business cycle effects and for monitoring the new self-employed after their entry to the market. Besides more detailed information on personal wealth, the panel also lacks a measure of risk aversion. Here, a time constant individual risk aversion is assumed to be part of the unobserved individual effect. However, an explicit analysis of the correlation between risk aversion and the transition into self-employment is not possible with the data set at hand.

Further research should include an analysis of the new entrepreneurs' success in the market, especially of those who became self-employed after the introduction of the EGZ. It would also be interesting to further investigate the influence of personal wealth on the transition into self-employment based on more accurate wealth measures.

Finally, a fixed effects Logit regression – if feasible – might be an interesting approach in order to avoid the distributional assumptions on the unobserved individual effects in the econometric model.

References

Arai, Bruce (1997): The Road Not Taken: The Transition from Unemployment to Self-Employment in Canada, 1961-1994. *Canadian Journal of Sociology*, (22/3), 365-382.

Audretsch, David, Martin Carree and Roy Thurik (2002): Does Entrepreneurship reduce Unemployment? Tinbergen Institute Discussion Paper TI2001-074/3.
http://www.tinbergen.nl/scripts/papers.pl?paper=01074.rdf (accessed in May 2008).

Baptista, Rui and Roy Thurik (2007): The relationship between entrepreneurship and unemployment: Is Portugal an outlier? *Technological Forecasting & Social Change*, (74), 75–89.

Blanchflower, David and Andrew Oswald (1998): What makes an entrepreneur? *Journal of Labor Economics*, (16/1), 26-60.

Caliendo, Marco, Alexander Kritikos and Frank Wießner (2006): Existenzgruendungsförderung in Deutschland. Zwischenergebnisse aus der Hartz-Evaluation. GfA Discussion Paper No. 05(2006).

Caliendo, Marco, Frank Fossen and Alexander Kritikos (2006): Risk Attitudes of Nascent Entrepreneurs: New Evidence from an Experimentally-Validated Survey. IZA Discussion Paper No. 2168.
http://www.iza.org/index_html?lang=en&mainframe=http%3A//www.iza.org/en/webcontent/publications/&topSelect=publications (accessed in May 2008).

Caliendo, Marco, Viktor Steiner and Hans-Joerg Baumgartner (2007): Kapitel IV – Mikroökonometrische Analyzen. In: Forschungsverbund IAB/DIW/sinus/GfA/infas: Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission, Arbeitspaket 1: Wirksamkeit der Instrumente, Modul 1e: Existenzgruendungen, BMAS, Berlin, S. 199-255.

Chamberlain, Gary (1980): Analysis of Covariance with Qualitative Data. *Review of Economic Studies*, 47, 225-238.

Chamberlain, Gary (1984): Panel Data. In: *Handbook of Econometrics*, Vol. 2, ed. Griliches, Zvi and Michael D. Intriligator. Amsterdam: North Holland, 1247-1318.

Constant, Amelie and Klaus Zimmermann (2004): Self-employment dynamics across Business Cycles: Migrants versus Natives. IZA Discussion Paper No. 1386.
http://ftp.iza.org/dp1386.pdf (accessed in July 2008).

Eichhorst, Werner and Klaus F. Zimmermann (2007): Dann waren's nur noch vier… Wie viele (und welche Maßnahmen der aktiven Arbeitsmarktpolitik brauchen wir noch? Eine Bilanz nach der Evaluation der Hartz-Reformen. IZA Discussion Paper No. 2605.
http://ftp.iza.org/dp2605.pdf (accessed in July 2008).

Evans, David and Boyan Jovanovic (1989): An Estimated Model of Entrepreneurial Choice under Liquidity Constraints. *The Journal of Political Economy*, (97/4), 808-827.

Evans, David and Linda Leighton (1989): Some Empirical Aspects of Entrepreneurship. *The American Economic Review*, (79/3), 519-535.

Evans, David and Linda Leighton (1990): Small Business Formation by Unemployed and Employed Workers. *Small Business Economics*, (2) 319-330.

Georgellis, Yannis, John G. Sessions and Nikolaos Tsitsianis (2005): Self-Employment Longitudinal Dynamics: A Review of the Literature. *Economic Issues*, (10/2). http://ssrn.com/abstract=921029 (accessed in July 2008).

Glaeser, Edward (2007): Entrepreneurship and the City. Harvard Institute of Economic Research Discussion Paper No. 2140. http://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID1001108_code347008.pdf?abstractid=1001108&mirid=1 (accessed in May 2008).

Glocker, Daniela and Victor Steiner (2007): Self-Employment: A Way to End Unemployment? Empirical Evidence from German Pseudo-Panel Data. IZA Discussion Paper Series No. 2561. http://ftp.iza.org/dp2561.pdf (accessed in July 2008).

Hansch, Esther (2006): Existenzgruendungen im Spiegel des Mikrozensus. *Wirtschaft und Statistik*, (5/2006), 495-504.

Holtz-Eakin, Douglas, David Joulfaian and Harvey Rosen (1994): Entrepreneurial Decisions and Liquidity Constraints. *RAND Journal of Economics*, (25/2), 334-347.

KfW Bankengruppe (2004): KfW Gruendungsmonitor 2004. http://www.kfw.de/DE_Home/Service/Online_Bibliothek/Research/PDF-Dokumente_Gruendungsmonitor/Gruendungsmonitor_2004.pdf (accessed in May 2008).

KfW Bankengruppe (2005): KfW Gruendungsmonitor 2005. http://www.kfw.de/DE_Home/Service/Online_Bibliothek/Research/PDF-Dokumente_Gruendungsmonitor/Gruendungsmonitor_2005.pdf (accessed in May 2008).

KfW Bankengruppe (2007): KfW Gruendungsmonitor 2007. http://www.ihk-kassel.de/solva_docs/kfw_gruendungsmonitor_2007_lang.pdf (accessed in May 2008).

KfW Bankengruppe (2008): KfW Gruendungsmonitor 2008. http://www.kfw.de/DE_Home/Service/Online_Bibliothek/Research/PDF-Dokumente_Gruendungsmonitor/KfW_Gruendungsmonitor_2008_lang.pdf (accessed in July 2008).

Knight, Frank H. (1964): Risk, uncertainty and profit. Kelley, New York, 1964 (reprint of the first edition from 1921).

Köllinger, Philipp and Christian Schade (2005): Warum Amerikaner mehr Unternehmen gründen als Deutsche. *Wochenbericht des DIW Berlin*, (45/2005), 673-680.

Kritikos, Alexander and Kathrin Kahle (2007): Kapitel I – Das Gruendungsgeschehen. In: Forschungsverbund IAB/DIW/sinus/GfA/infas: Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission, Arbeitspaket 1: Wirksamkeit der Instrumente, Modul 1e: Existenzgruendungen, BMAS, Berlin, S. 40-90.

Marlow, Susan and David Storey (1992): New Firm Formation and Unemployment: A note on Research Method. *International Small Business Journal*, (10/3), 62-67.

Meager, Nigel (1992): Does Unemployment lead to Self-Employment? *Small Business Economics*, (4), 87-103.

Meyer, Bruce (1990): Why are there so few Black Entrepreneurs? NBER Working Paper No. 3537.

Mundlak, Yair (1978): On the Pooling of Time Series and Cross Section Data, *Econometrica*, 46, 69-85.

OECD (1986): Self-Employment in OECD Countries, in: OECD Employment Outlook, 1986.

Petrova, Kameliia (2005): Part-Time Entrepreneurship and Wealth
Effects: New Evidence from the Panel Study of Entrepreneurial Dynamics. *Working Paper*. http://129.3.20.41/eps/mic/papers/0510/0510006.pdf (accessed in May 2008).

Pfeiffer, Friedhelm (1994): Selbständige und abhängige Erwerbstätigkeit. Arbeitsmarkt- und industrieökonomische Perspektiven. Frankfurt a. M., 1994.

Pfeiffer, Friedhelm (1999): Existenzgruendungpotentiale unter Arbeitsuchenden: Empirische Evidenz auf der Basis des Mikrozensus. ZEW Discussion Paper No. 99-02.

Pfeiffer, Friedhelm and Frank Reize (2000): Business start-ups by the unemployed – an econometric analysis based on firm data. *Labour Economics*, (7), 629–663.

Ritsilä, Jari and Hannu Tervo (2002): Effects of Unemployment on New Firm Formation: Micro-Level Panel Data Evidence from Finland. *Small Business Economics*, (19) 31–40.

Schmidt, Simone (2000): Erwerbstätigkeit im Mikrozensus.
Konzepte, Definition, Umsetzung. ZUMA-Arbeitsbericht 2000/01, ISSN 1437-4110. http://www.gesis.org/Publikationen/Berichte/ZUMA_Arbeitsberichte/00/00_01.pdf (accessed in July 2008).

Schuetze, Herb (2000): Taxes, economic conditions and recent trends in male self-employment: a Canada–US comparison. *Labour Economics*, (7) 507–544.

Schumpeter, Joseph A. (1946): Kapitalismus, Sozialismus und Demokratie. Francke, Bern, 1946.

Storey, David (1991): The Birth of New Firms – Does Unemployment Matter? A Review of the Evidence. *Small Business Economics*, (3), 167-178.

Thurik, Roy and Ingrid Verheul (2002): The Relationship between Entrepreneurship and Unemployment: The Case of Spain, *Working Paper*, http://www.few.eur.nl/few/people/verheul/hvi_6.pdf (accessed in June 2008).

Thurik, Roy et al (2007): Does Self-Employment reduce Unemployment? Discussion Papers on Entrepreneurship, Growth and Public Policy No. 705.

Van Praag, C.M. and J.S. Cramer (2001): The Roots of Entrepreneurship and Labour Demand: Individual Ability and Low Risk Aversion. *Economica*, (2001), 45-62.

Wagner, Joachim (2002): The Impact of Risk Aversion, Role Models, and the Regional Milieu on the Transition from Unemployment to Self-Employment: Empirical Evidence for Germany. IZA Discussion Paper No. 468.
http://ftp.iza.org/dp468.pdf (accessed in July 2008).

Wagner, Joachim (2007): What a Difference a Y makes-Female and Male Nascent Entrepreneurs in Germany. *Small Business Economics*, (28), 1-21.

Wießner, Frank (1998): Das Überbrückungsgeld als Instrument der Arbeitsmarktpolitik – eine Zwischenbilanz. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, (31/1), 123-142.

Wooldridge, Jeffrey M. (2002): Econometric Analysis of Cross Section and Panel Data. MIT Press, London, 2002.

Curriculum Vitae

Name	Veronika Regina Scharfenberg
	geb. Schreiner
Date/Place of Birth	12/31/1981, Munich
Nationality	German

Education

10/2005-09/2008	LMU München, Munich Graduate School of Economics
	Ph.D in Economics
2001-2005	Universität Passau
	Diplom-Kauffrau
2000-2001	Universität Passau
	Undergraduate cultural studies
2000	Gymnasium Vaterstetten, Baldham
	Abitur

Professional Experience

10/2008 - current	Consultant, d-fine GmbH
05/2006-07/2007	Teaching Assistant, Seminar für Empirische Wirtschaftsforschung, LMU München
03/2004-04/2004	Intern, TNS Infratest München
08/2003-10/2003	Intern, T-Systems, Berlin
08/2002-09/2002	Intern, Lloyd Textile, Hong Kong
04/2002-07/2002	Tutor, Lehrstuhl für Statistik, Universität Passau
09/2001	Intern, BMW, München

Scholarships

10/2005-09/2008	Deutschen Forschungsgemeinschaft: Ph.D. scholarship
03/2003-07/2003	DAAD: scholarship for studies abroad
10/2000-03/2005	Scholarship for the very gifted (BayBFG)

Languages	German, English, French, Spanish, Latin