Three Essays in Empirical Economics

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

2007

vorgelegt von
JOY KAPFER

Referent: Professor Dr. Joachim Winter
Korreferent: Professor Dr. Ludger Wößmann

Promotionsabschlussberatung: 6. Februar 2008

for my beloved husband Josef

and

for my beautiful beautiful babies – Luca & Eliana

Contents

List of Figures

Chapter One

Chapter Two

Chapter Three

List of Tables

Chapter One:

Chapter Two

Chapter Three

Acknowledgements

I am in great debt to my supervisor, Professor Joachim Winter, for his continuous support throughout these three essays. He has been patient and encouraging from the very beginning stage of my thesis. He inspired and disciplined me. He encouraged me to explore the world of empirical economics. Without his academic guidance and instructions, it would be impossible to accomplish my goals.

I am also very grateful to Professor Ludger Wößmann and Professor Flaig Gebhard for willing to be the second and third thesis advisors on my committee.

I would like to thank the Manpower Demonstration Research Corporation in New York, USA for providing the data file used in my first essay, which studies the effect of the Florida's Family Transition Program on reducing unemployment. I benefited from Dr. Joachim Wolff for sharing his STATA program for generating the graphics for hazard rates in my first essay. My thanks also go to Deutsches Institut für Wirtschaftsforschung, DIW Berlin (German Institute for Economic Research Berlin) for kindly providing the German Socio-Economic Panel datasets, which was used in my second and third essays to investigate the life insurance and supplementary health insurance market. I would also like to thank Markus Grabka from DIW for sharing the computer program used in my second essay to generate the relevant tax variables, and Brian Poi from STATA Corporation for answering my STATA question about a two-equation system in my third essay, and Dieter Riefle from Versicherungs Kammer Bayern for providing a lot of useful background information regarding life insurance and supplementary health insurance.

My research work benefited from helpful comments and suggestions from the participants at Research Strategy Seminar and Research Seminar in Munich and the doctoral conference in Dortmund. Especially, I would like to thank Betty Tao, my former colleague in USA and my good friend, for proofreading my English and providing her advice on my thesis work.

I am benefited quite a lot from the Ph.D courses I took at the economics department of University of North Carolina at Chapel Hill, USA before I started my thesis in Munich. I gratefully acknowledge the financial support from Deutschen Forschungsgemeinschaft (DFG) from 2004 to 2007, as well as the scholarship from University of North Carolina at Chapel Hill from 2002 to 2004.

Finally, I would like to thank the members of my family for their support. I dedicate my thesis work to my husband Josef, who shared my difficult moments and gave me support and encouragement the whole time, and to my lovely babies Luca and Eliana, who motivated me to work hard while they slept. As this three-year race draws to an end, I am deeply proud, not only of my thesis, but also of my two angels.

Preface

Real empirical work is question-driven. This is what I learned from doctoral lectures in graduate school when I was in U.S.A. The motivation of doing empirical work can be many-fold: evaluating public policy, testing an economic theory, estimating a parameter of interest, understanding the micro-behaviour of agents or macro-behaviour of an economy, etc.

The primary focus of this thesis is to explore socially and economically relevant questions. It contains three independent essays in empirical economics. Chapter One investigates single mothers' labour force participation from a random experiment in U.S.A. Chapter Two examines the social factors behind the holding of life insurance in Germany. Chapter Three studies the effect of self-reported health on the purchase of supplementary health insurance in Germany. All three essays deal intensely with empirical datasets and econometric models.

Chapter One is my starting point in empirical work, which piqued my interest in empirical economics. After reading a paper by Grogger and Michalopoulos (2003), which studied the effect of time limits on welfare use with data from Florida's Family Transition Program, a thought jumped into my mind that there might be more to be explored with this randomized dataset. I started to study the background of Florida's Family Transition Program and tried to find a research topic out of it. The talks with my thesis advisor gave me the idea to examine single mothers' labour force participation under the program effect. Evaluation techniques are applied to identify the effect. I analyze the probability of being employed and then estimate the conditional probability of exit from unemployment into different employment states with a discrete time proportional hazard model. The main finding is that, though the program has generated certain gains in employment in the middle of the study period, these gains diminished at the end of the follow-up period. The results from the hazard model indicated that

the program effect on exit from unemployment is also not significant after controlling for observed characteristics and unobserved heterogeneity.

Chapter Two focuses on the social factors behind holdings of life insurance in Germany. Since I lived in Germany, my husband and I paid a monthly life insurance premium to a private health insurance company. As I realized that life insurance is an important part of private household saving in Germany, this topic became important to me. With a nation-wide representative dataset from German Social Economic Panel Studies, I examined the consequences of change of family structure, change in labour force participation and change in tax code on life insurance holdings. Chapter Two shows that the bequest motives are quite strong among families with dependent children, being unemployment has significant negative effect and higher tax code is related to higher probability of life insurance holdings due to incentives from German tax system. Since tax benefits of life insurance was abolished at the end of 2004, I evaluate the program anticipation effect using the same dataset and it shows that people responded positively to the program change with a higher probability of purchase.

Chapter Three studies risk selection in private health insurance market in providing supplementary contracts. This study was inspired by some casual talks with people working in the insurance business about patients' data protection. I learned a story about one person who hid information about his previous in-patient treatment, causing his insurance company to cancel the supplementary contract with him. I started working on this research project with the data from German Social Economic Panel Studies to find out whether there exists risk selection in private health insurance market for supplementary coverage. The main explanatory variable is self-reported health status. After controlling for measurement error in this variable, the results show that bad health is related to a lower probability of purchasing a supplementary contract, other things being equal. The health insurance companies, through its risk selection procedure, insure only good risks and exclude bad risks. The evidence of risk selection provides some guidelines for policy discussions about patients' data protection.

While I was working on these three essays, I was brought into a world of empirical econometrics. From the discrete time proportional hazard model in the first essay to a panel probit model correcting for auto-correlated errors in the second essay, and to a two-equation system controlling for measurement error and a Heckman selection model in the third essay, I am amazed at the fact that econometric models can be such a powerful tool for empirical research. Yet there are more to be learned and there are more empirical questions to be answered. This thesis is my start.

Chapter One

From Welfare to Work: Has Florida's Family Transition Program Reduced Unemployment?

This chapter studies the effect of Florida's Family Transition Program (FTP) on single mothers' labour force participation. As a pilot program implemented 2 years prior to establishment of national wide TANF welfare program that is currently in operation, the FTP program provided an important test case of the new law. I use the 4-year administrative data from the FTP program and find that the FTP program had generated some impact on the labour force participation, especially on being full-time employed. However, the employment gains are mainly concentrated in the middle of the study period. Moreover, the estimation from a discrete time proportional hazard model with unobserved heterogeneity shows that the FTP program is statistically insignificant in speeding up the exit from unemployment to take the first available job since the random assignment.

I: Introduction:

Over the years, the objectives of welfare reform have been to reduce dependency and promote work while still alleviating needs. In 1996, the passage of the law "The Personal Responsibility and Work Opportunity Reconciliation Act" (PRWORA) replaced AFDC (Aid to Families with Dependent Children) with TANF (Temporary Aid to Needy Families). The establishment of RRWORA has been seen as part of the federal effort to 'end welfare as we know it' under Clinton Administration. These reforms are significant in American welfare policy as they changed benefit structure, introduced time limits on the receipt of benefits, strengthened requirements for mandatory participation in work-related activities, and changed various administrative procedures.

Not only have important changes in welfare programs been observed in the last decade, but also remarkable welfare-related outcomes. According to a synthesis conducted by RAND, the welfare caseload fell to approximately 2.1 million families as of September 2001, less than half its all-time peak level of 5.0 million families in 1994; the fraction of welfare recipients participating in welfare-to-work activities or working increased rapidly; the employment rates and earnings of single mothers rose substantially; family income increased and the poverty rates declined. Meanwhile, the upward trend of non-marital fertility rate has levelled off.

There has been great research effort devoted to understanding the effect of TANF prior to and after its implementation. The effects of welfare reform as a whole have been found to substantially reduce the welfare caseload and it seems likely that welfare reform is responsible for a portion of an increase in work and earnings among single mothers during the last decade. However, the studies of the depth of the impact are confounded by contemporaneous policy changes such as increases in the Earned Income Tax Credit (EITC), expansion of subsidized health insurance de-linked from welfare receipt, and increases in minimum wages that took place during the same time, and the long robust economics expansion in last decade. The imposition of time limits by TANF is believed by many researchers to have exerted great impact on the reduction of welfare caseloads and thus it

became the most broadly studied subject. The effect of time limits has been considered in research such as the Council of Economic Advisers (1997, 1999), Swann (1998), Moffitt (1999), and Ziliak et al. (2000). The study of the effects of time limits is not only confounded by the contemporaneous policy changes and the favourable economic growth as mentioned above but also the other policy components of TANF such as financial incentives, mandatory work-related activities and parental responsibilities.

This chapter explores the data from a randomized experiment, Florida's Family Transition Program (FTP), that is conducted in Escambia county in Florida to study the overall program impact on the labour force participation of single mothers. Some previous studies of the FTP program include Bloom et al. (2000), Grogger and Michalopoulos (1999, 2003), etc.

The empirical analysis in this chapter is composed of two parts: the analysis of probability of being employed and the estimation of the hazard rate of exit from unemployment. For the first part I classify the employment status according to their quarterly earning statistics as follows: irregularly part-time employment, part-time employment, full-time employment[1]. I then examine the employment probability over a 16-quarter period after the random assignment. For the second part, a discrete time proportional hazard model is applied to study the conditional probability of exit from unemployment into three different employment states.

The FTP program has generated certain gains in employment in the middle of the studying period, especially in part-time and full-time employment. For part-time employment, there is an apparent and consistent gap up to the 10th quarter between these two groups but at the end the gap becomes indistinguishable afterwards. The probabilities of full-time employment are increasing for both groups over the 16 quarters and the differences between these two groups become apparent and stable after the 2nd quarter. However, at the end of the follow-up period, the difference diminished and the AFDC group caught up. FTP's results

[1] Employment status classification I used is different from Bloom et al. (2000) in which they simply defines the outcome variable, employment, as non-zero earnings.

were affected by the unusual environment in which it was operated — a period of low unemployment, highly publicized changes in state and national welfare policies, and an unprecedented 70 percent decline in Florida's welfare caseload. The expectation of the AFDC group was that they would join the new program after the study period motivated them to somehow be more active in the labour market participation, which helps to illustrate why there was little room left for FTP to generate large impacts at the end of the study period.

The results from the hazard model indicate that the program effect on exit from unemployment, after controlling for observed characteristics and unobserved heterogeneity, is also not significant. Assigning the single mothers who were being unemployed to the FTP program did not accelerate their job take-up rates. As the hazard model studies the single unemployment spell during which the random assignment happened, the fact that the single mothers exited and entered unemployment more than once is not considered. The insignificant program impact on picking up a first job after random assignment is not unexpected. As the FTP program offered some valuable services such as job training, the single mothers assigned to the program would find it more interesting to stay in the program and take advantage of these services in the short term, regardless of the financial incentives and mandatory working-hour requirements.

Section II of this chapter provides background on the FTP data and some descriptive statistics. Section III discussed how I use the information from the original dataset to re-define the employment status. Section IV analyzes the employment probabilities over the 16 quarters after the random assignment. Section V presents a discrete time proportional hazard model with unobserved heterogeneity and the estimates. Section VI concludes.

II. Data

A. Background

Florida's Family Transition Program (FTP) was created by the Florida legislature and provides a useful tool to study the confounded effect of the welfare reform. It operated from 1994 through 1999 as a pilot program in Escambia

County (Pensacola) in Florida under waivers of pre-1996 welfare rules. It anticipated key elements of the federal law-- time limits on cash assistance receipt, financial incentives, together with an array of services and supports. For this reason — and because Escambia was the first place in the U.S where people reached a time limit and actually had their benefits cancelled — FTP was an important test case for states and localities across the country[2].

Starting from May 1994, all welfare recipients in Escambia County were randomized into one of two groups: the treatment group or the control group[3]. New entrants were randomized at the time they applied for benefits while ongoing recipients were randomized at the time of their semi-annual re-certification interviews. The treatment group was subjected to FTP policy reforms that include time limits, financial work incentives, and mandatory work-related activities. The control group was enrolled under the traditional AFDC program, where welfare benefits were an entitlement so that all poor single-parent families with at least one child under 18 years of age were eligible to receive aid[4]. People were assigned to the FTP and AFDC groups from May 1994 through October1996[5].

The policy components in FTP and AFDC can be summarized in Appendix 1. The three main categories of reforms from Appendix 1 are: time limits, financial work incentives, and enhanced welfare-to-work services. FTP families faced a 24-month time limit in any 60-month period. Particularly disadvantaged families (about 53 percent of the sample) received a 36-month time

[2] However, Florida's current statewide welfare program includes similar time limits and financial work incentives, but differs from FTP in other key respects[2]; thus, the evaluation is not necessarily assessing the state's current program.

[3] The randomization process was conducted by a computer program. It does not depend on the asset information. Exemptions from random assignment are given to those who are disabled, or older than 62 years old, or under 18 years old school attendants.

4. Although the Work and Gain Economic Self-Sufficiency (WAGES) program replaced AFDC state wide in Florida in late 1996, to facilitate completing the study, both FTP and traditional AFDC continued to operate in Escambia County until late 1999. FTP officially ended on December 1, 1999, when individuals in the FTP and AFDC groups became subject to WAGES rules. However, distinction between the groups began to blur in September 1999, when AFDC group members were informed that they would become subject to WAGES in December.

5. Beginning in October 1996, new applicants for welfare in Escambia County who had not already been assigned to the FTP group or the AFDC group were placed into WAGES.

limit in any 72-month period[6]. After that, their benefit could be cancelled. Overall, about 21 percent of the FTP group received at least as many months of benefit as their time limit allowed and the others left welfare before reaching their time limit. AFDC families faced no time limits. They were entitled to benefit receipt as long as they have a child under 18 years old.

FTP families enjoyed relatively generous financial incentives as well. The first $200 of monthly earnings was disregarded from income in determining their monthly benefits, and earnings in excess of $200 were subject to a benefit reduction rate, or tax rate, of 50 percent. ADFC families faced the conventional financial incentives, which were less generous. The disregarded amount of earning not only is less than for FTP families but also decreases with the length of time. The benefit reduction rate is 100 percent. FTP participants received subsidized transitional childcare for two years after leaving welfare for work, as opposed to the one year provided under prior rules. These financial incentives clearly provided FTP families motivation from welfare to work.

Both FTP and AFDC families were required to spend 30 hours per week in mandatory work-related activities. The exemptions applied only to mothers with infants under 6 months of age under FTP. Under AFDC, the exemptions applied to mothers with children under 3 years old. FTP participants were also more likely than AFDC group members to be required to participate in employment-related activities, and the program developed some enhanced education, training, and job placement assistance services. To the extent that FTP's enhanced services were valuable and were tied to welfare, some FTP families might postpone the working decision to accumulate more human capitals.

The data analyzed in this chapter comes mainly from the combined data file of administrative records and a short survey instrument known as the Background Information Form. The administrative records tracks the welfare receipt, earning status, employment status, Food Stamp receipt, Unemployment

[6] Recipients were limited to 36 months of welfare in any 72-month period if they (1) had received AFDC for at least 36 of the 60 months prior to enrollment or (2) were under 24 years old and had no high school diploma and no recent work experience.

Insurance receipt up to 10 quarters prior to the random assignment and up to 20 quarters since the random assignment. The survey was administrated to welfare applicants and those who are re-certified at the time when they applied for welfare receipt or were re-certified. It contains basic demographic characteristics, such as education, marital status, age group, ages of all kids in the family, working experience, etc. Due to the confidentiality issue, the relevant calendar information is dropped out of the public use file.

B. Descriptive statistics

Our sample includes the 2810 single-parent families assigned to the random experiment from May 1994 to October1996. Since I focused on single mothers, I dropped male observations. After deleting the observations with missing values, I have 2657 observations left in the dataset, 1336 of those were assigned to the FTP group and 1321 to the AFDC group. To check the randomization is properly executed, I test the equality of means of different variables between these two groups.

Table 1.1 summarizes the basic demographic characteristics for both FTP and AFDC groups. Overall, these demographic characteristics are quite similar between these two groups, with the p-values for testing equality of means are all above the level of 5 percent. Therefore the initial randomization was properly executed. Like welfare recipients elsewhere, the sample members are disproportionately black (around 52 percent) and low educated with 11 years of schooling on average[7]. Their low education status likely presented them a great barrier in job market. In addition, they had weak labour market attachments and high levels of welfare usage. At the time of random assignment, the ongoing recipients and new entrants were assigned to the two groups randomly so that there was no systematic behavioural difference. Moreover, around 45 percent members of both groups were falling into the age category of between 25 and 34, which means nearly half of the observations were at the prime working age.

[7] For the highest degree obtained, around 44 percent of the sample members have a high school diploma, and around 40 percent have no degree at all.

Implementing a program that offered proper work incentives thus had important implications in terms of self-efficiency and anti-poverty.

Table 1.1 --Summary Characteristics of The FTP and AFDC Groups at The Time of Random Assignment

Characteristics	**FTP Group (μ_1)**	**AFDC Group (μ_0)**	**Mean Difference (μ_1-μ_0)**	**P-Value (Ho:μ_1-μ_0=0)**
Mother black	52.4	52.0	0.4	0.84
Years of schooling	11.0	11.1	-0.1	0.06
Number of kids	2.0	2.0	0.0	0.59
Age of youngest kids	5.0	5.2	-0.2	0.40
Month of welfare utilization in 24 months prior to random assignment	12.4	12.6	-0.26	0.49
Quarters of employment in 2 years prior to random assignment [8]	2.1	2.1	0.0	0.64
New entrants	51.9	49.8	2.1	0.28
Mother age less than 20	8.2	6.4	1.8	0.49
Mother age between 20 and 24	25.2	26.0	-0.7	
Mother age between 25 and 34	44.6	44.7	-0.1	
Mother age between 35 and 44	18.7	19.8	-1.0	
Mother age greater than 45	3.1	3.1	0.0	
Number of observations	*1336*	*1321*	--	--

Note: Tables entries are sample means; for the age categories of single mothers, Pearson Chi2 test is applied.

In Table 1.2, I summarize the working experience in further detail for the FTP and AFDC group. Because of the randomization there is no significant difference between these two groups in term of past working history or job searching behaviour. In the Private Opinion Survey, single mothers have reported several reasons, such as arranging child care, family problems, health or emotional problems, etc., for not being able to work. Overall, the single mothers

[8] For simplicity reason, here the dummy of employment status takes the value of 1 as long as the corresponding quarterly earning is greater than 0.

in the sample lacked recent labour market experience, which also explains to some extent the low job take-up rates in my hazard model. However, the original dataset didn't provide important information about the past job such as economic sector, profession, distance to work, etc. Without this information, it is difficult to estimate the individual's decision to take a specific job offer after the random assignment.

Table 1.2 -- Comparison of Working Experience Features of The FTP and AFDC Groups at The Time of Random Assignment

Variables	FTP Group (μ_1)	AFDC Group (μ_0)	Mean Difference (μ_1-μ_0)	P-Value (Ho:μ_1-μ_0=0)
Never worked before random assignment	9.3	9.7	-0.4	0.72
Ever worked full-time 6 months or more	58.9	59.3	-0.4	0.82
Looked for work in the past 12 months	35.2	33.4	1.8	0.33
Employed in the past 12 months	46.1	44.7	1.4	0.45
Currently employed	16.1	17.0	-0.8	0.59

Note: Tables entries are sample means.

Table 1.3 summarizes the welfare dependency for both groups. I define the short-term recipient as staying on welfare less than 2 years and long-term recipient longer than 2 years. It is worth noting that on average the sample members are quite dependent on welfare, with 52 percent of the FTP and 56 percent of the AFDC being long-term recipients. For them, cash assistance from the welfare program was one of the most important resources for their families. The FTP program, which promoted self-sufficiency through work, limited benefit receipts to a much shorter period. Considering their job barriers, one result of the

FTP program might be that it would put these families that relied on welfare to a large extent into an economically worse situation[9].

Table 1.3 -- Comparison of Welfare Dependency of The FTP and AFDC Groups at The Time of Random Assignment

Variables		FTP Group (μ_1)	AFDC Group (μ_0)	Mean Difference (μ_1-μ_0)	P-Value (Ho:μ_1-μ_0=0)
Short term recipient (less than 2 years)		36.0	33.1	2.9	0.11
Long term recipient (greater than 2 years)		52.2	55.9	-3.7	0.06
Never receiving AFDC		11.8	11.0	0.7	0.53
Level of economically disadvantaged[10]	1	16.6	16.1	0.5	0.79
	2	64.8	66.0	-1.3	
	3	18.7	17.9	0.8	

Note: Tables entries are sample means; for the 3 levels of economically disadvantages, Pearson Chi2 test is applied.

Comparing the mean statistics of the FTP and AFDC groups, the pre-treatment characteristics are not systematically different. Since the randomization is well conducted, it provides a rather convenient approach for the analysis of the program impact on the employment probability in next sections. I can therefore assess the effectiveness of the program without further identification issues since the effects of the external economical condition and other programs such as food stamp, social security support on both groups are similar.

III: Employment Probabilities

A. Redefining the Employment Status

[9] The MDRC evaluation report indicates that 'FTP did not affect hardships associated with material well-being, food security and the need to rely on social services'.

[10] Level 1 is defined as on welfare greater than 2 years, no high school diploma, no work in past year; level 3 is defined as not on welfare 2 or more years, has worked in past year, has high school diploma; level 2 is everybody else.

I set the analysis period to 8 quarters prior to and 16 quarters after the random assignment[11]. The quarterly employment dummy in the original data is defined as long as earnings are greater than zero, which is too general to reveal more detailed information about part-time or full-time employment[12]. I therefore create the new employment status according to earnings statistics and define it as follows: irregularly part-time employed; part-time or irregularly full-time employed; regularly full-time employed.

Figure 1.1: Mean Earnings for FTP and AFDC Groups

(Prior to Random Assignment)

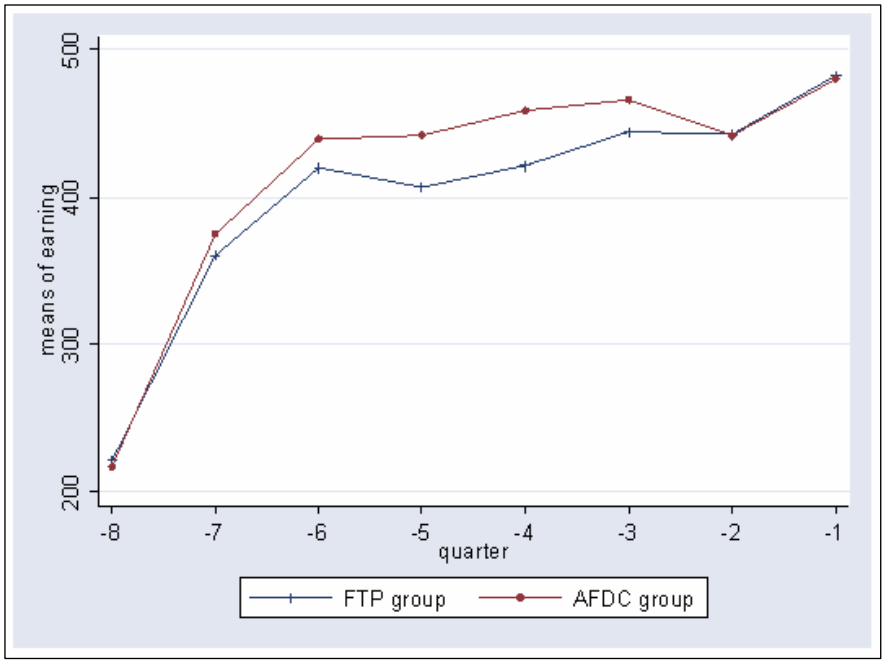


Figure 1.1 plots the mean earnings of the FTP and AFDC groups 8 quarters prior to the random assignment. As it shows, the average earnings for both groups were quite low. The AFDC group had relatively higher average quarterly earning than the FTP group, though the difference is not so high (with

[11] In the original dataset, the employment status was recorded up to 10 quarters prior to and 20 quarters after the random assignment. Since not all the members were well followed, there are a lot of missing values from 9th quarter prior to and 17th quarter after.

[12] This employment definition is also adopted by Bloom at et all (2000).

the largest one amount to $50). From the second quarter prior to the random assignment on, these two groups had quite similar average quarterly earnings, which also reinforced the randomization of eligibility when the experiment happened. The median of the quarterly average earnings are all zero, which implies that at least 50 percent of both groups had no earnings from work at all. This is no doubt related to the low employment rate for these single mothers.

Figure 1.2: Mean Earnings for FTP and AFDC Groups

(After Random Assignment)

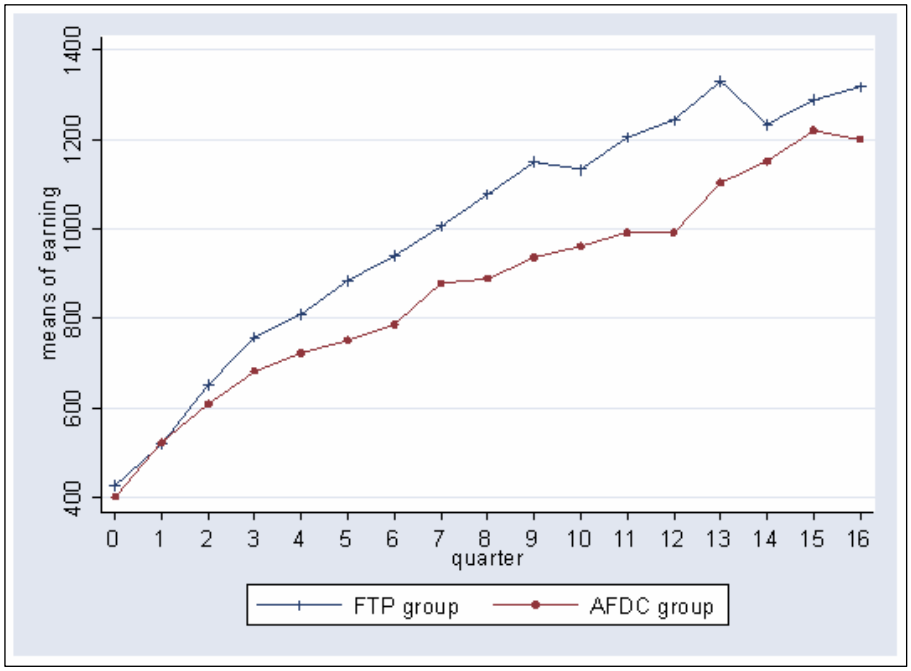


Figure 1.2 plots the means of quarterly earnings after the random assignment. Quarter 0 is when the random assignment happened. As it shows, the average earnings between these two groups at the first quarter after random assignment were not very different from each other. However, starting from quarter 2 after the random assignment, the gap between these two groups started becoming positive and increased rapidly in year 2 and 3. In quarter 4, the average earning of the FTP members already exceeded that of AFDC members by around 12%, in quarter 8 by 21%. The incentive to work and earn money appeared

stronger in the FTP group. Starting from quarter 14, the gap reduced and on average, the FTP members earned only 7% more than their counterparts and 10% more at the end of the follow-up period. One explanation would lie in the state wide implementation of WAGES (Work and Gain Economic Self-Sufficiency) after year 1996. Even though those who remained in Escambia County were not actually subject to WAGES until after the study ended, the behaviour of some AFDC group members, exposed to the heavy publicity, was affected to some extent. In addition to the favourable macroeconomic environment where there was steady and stable economic growth during mid 1990's, this can somehow explain the upward trend in average earnings across time. Unfortunately, I do not have a longer follow-up period. The trend of the earnings pattern is thus unknown. The FTP group might have only a shorter period gain over the AFDC group. Over a longer horizon, the gain might disappear and these two groups might finally have the same average earnings.

As the sample members I study here faced more serious work barriers than other groups in the society, I define the part-time and full-time employment in a relative loose way. First I define the potential quarterly earnings for part-time and full-time employment. I take the minimum wage of $4.25 per hour in year 1994 in Florida as the potential wage the single mother would earn if she took a job that is mainly low paid due to her limited working experience and education. Part-time working hours per week are defined as a minimum of 10 hours. Full-time working hours per week are defined as 35 hours. Therefore, I could calculate the potential quarterly earning as $510 for the part-time employment and $1785 for the full-time employment. Therefore, if a sample member has zero earnings for a specific quarter, then her status is unemployed. If her earnings are greater than zero but less than $510, then she is classified as irregularly part-time employed for that quarter. If her earnings are greater than $510 but less than $1785, then she is classified as part-time employed or irregularly full-time employed and as full-time employed if earnings are more than $1785.

B. Probabilities of Employment of Three Different Forms

Figure 1.3 plots the probabilities of irregularly part-time employment for the FTP and AFDC group. Overall, both groups have shown downward trends across time since the random assignment. However, there was no significant difference between these two groups. The test of equality of means yields a high p-value of 41 percent; therefore I can not reject the hypothesis that both groups had similar means.

Figure 1.3: Probability of Irregularly Part Time Employed

(After Random Assignment)

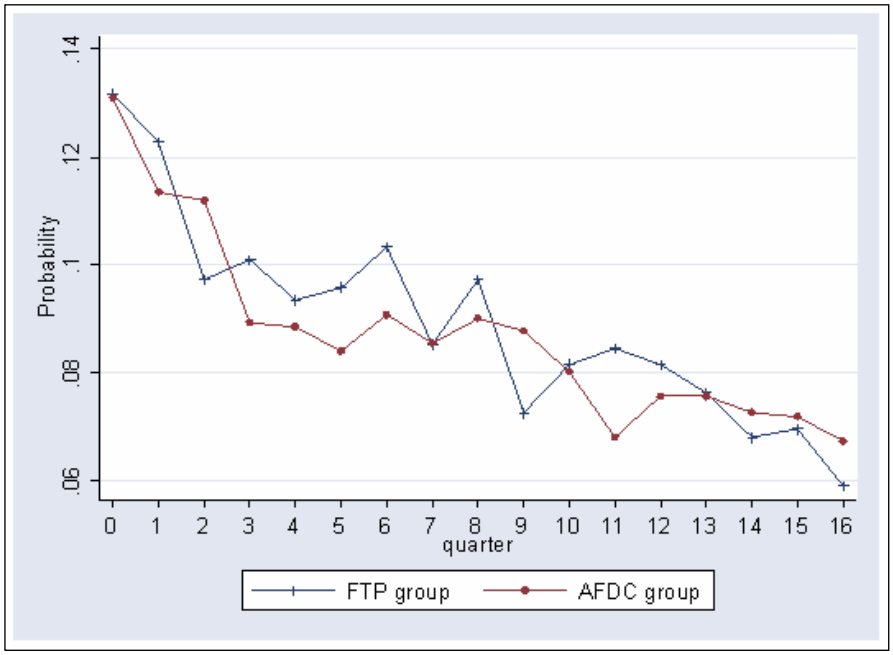


Figure 1.4 plots the probabilities of part-time employment for the FTP and AFDC group. During the first 12 quarters after random assignment, the probability for the FTP group was consistently higher than for the AFDC group and it peaked in the 8th quarter with 17 percent, with 4 percent more than the AFDC group. However, the last 4 quarters showed a reverse situation, with a lower probability of being part-time employed for the FTP group than for the AFDC group. Over the four years follow-up period, the impact of the FTP program on taking part-time job increased in the first two years and then decreased dramatically in the next 2 years. A test of equality of means yields a p-

value of 0.1 percent, which is highly significant and strongly rejects the hypothesis that both groups had the same means.

Figure 1.4: Probability of Part Time Employed

(After Random Assignment)

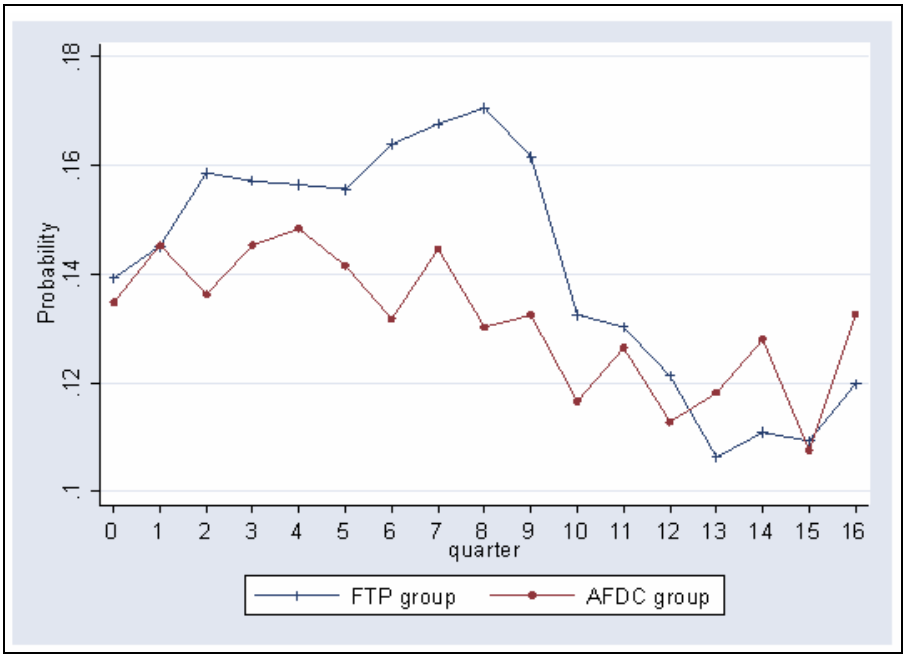


Figure 1.5: Probability of Full Time Employed

(After Random Assignment)

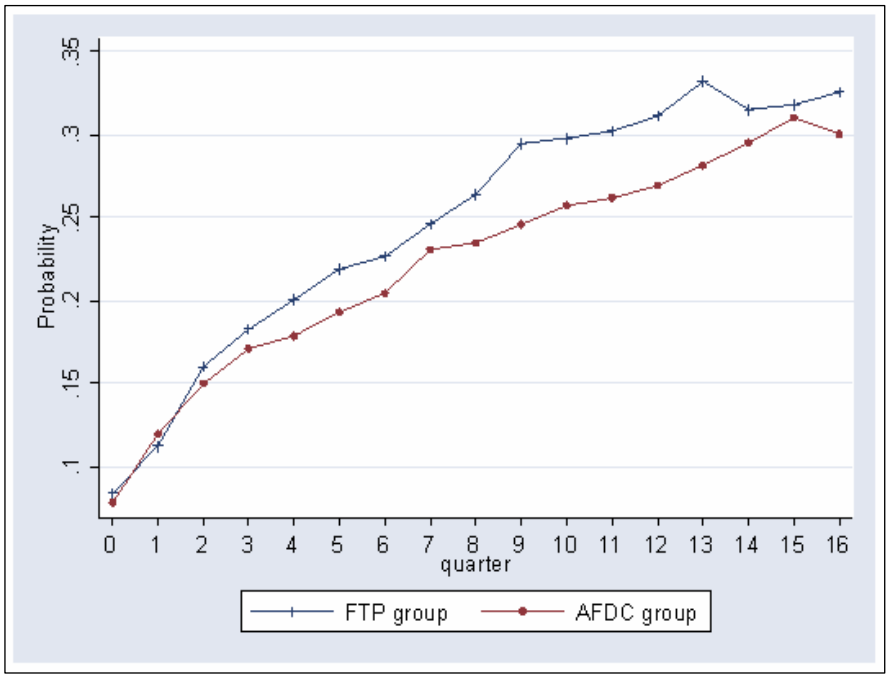


Figure 1.5 plots the probability of full-time employment for the FTP and AFDC group. Except in the first quarter after random assignment when the probability was slightly lower for the FTP group than for the AFDC group, it was consistently higher through the follow-up period. Starting from around 8 percent at the time of the random assignment for both groups, both trends were upward increasing, with 33 percent for the FTP group and 30 percent for the AFDC group in the last quarter. Apparently, there was an upward time trend that showed that the FTP program was operated under a favourable economic environment where the overall employment was steadily increasing. The gap between the FTP and AFDC group was increasing gradually across the time and peaked at the third year but then decreased in the fourth year. The test of equality of means yields a high p-value less than 0.1 percent. Therefore the difference in probabilities of full-time employed was statistically higher for the FTP group than for the AFDC group. Moreover, Figure 1.5 tracks quite closely the earnings pattern compared to Figure 1.2. Gains in earnings were mostly contributed by the full-time employment.

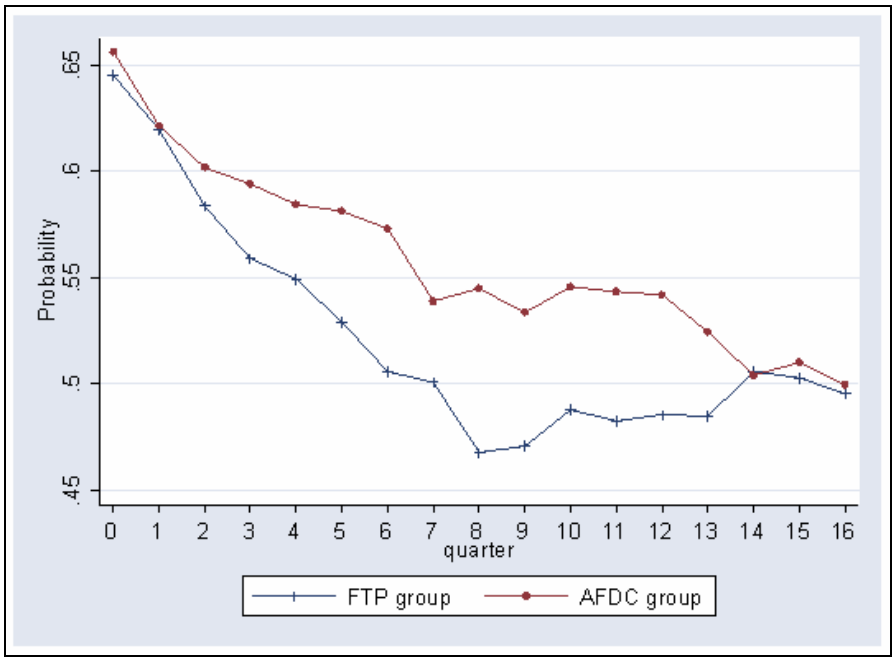
Figure 1.6 shows the probabilities of unemployment. It is obvious that the FTP group had lower probabilities across all the quarters after the random assignment. Especially in year 2 and 3, the difference in probability was larger than in year 1 and year 4. The FTP program had significant impact on reducing unemployment. Actually in the later period, the probability of being unemployed was increasing for the FTP group and the difference between these two groups were diminishing. Even though the unemployment probability was decreasing with the time for both groups, it is still noticeable that the probabilities were quite high. At the time of random assignment, they were around 65 percent. Then they were reduced to around 50 percent in the last quarter. The overall downward trend can be attributed the reasons like the favorable economic condition in mid 1990s and the anticipation effect of participating in WAGES after 1996.

Overall, FTP had a positive effect on being part-time and full-time employed. In this sense, FTP has fulfilled its target of encouraging the single mothers to seek means for self-sufficiency. However, the program had effects only in the very short term, about 2 or 3 years after the implementation. At the end

of the follow-up period, the difference diminished and the AFDC group caught up. The unusual environment under which the FTP program was operated — a period of low unemployment and stable economic growth, highly publicized changes in state and national welfare policies shaped the outcomes of the AFDC group, and therefore left little room for FTP to generate significant impacts.

Figure 1.6: Probability of Unemployed

(After Random Assignment)



C. Employment Probability Based on Different Characteristics

To examine further how the probabilities of employment differ amongst the different characteristics of the FTP and AFDC groups, I group the observations at the time of the random assignment into short term and long term recipients, new entrants and ongoing recipients, first time and non-first time applicants, families with and without preschool-age children, and the high and

low level of economically disadvantaged. Within each group, I examine two types of employment possibilities: part-time employment, and full-time employment[13].

Figure 1.7: Employment Probability for Short-Term Recipients

(After Random Assignment)

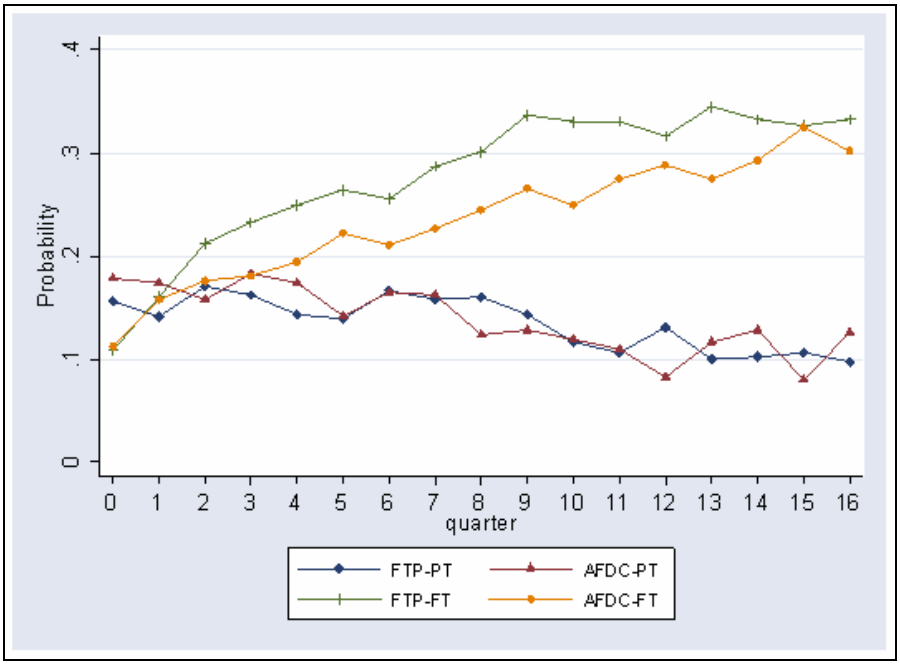


Figure 1.7 first plots the employment probabilities for the short-term recipients staying on welfare for less than 2 years at the time of the random assignment, while Figure 1.8 plots the employment probabilities for the long-term recipients staying on welfare for longer than 2 years. Comparing these two graphs provides some interesting interpretations. First of all, at the time of random assignment, the probabilities of being part-time or full-time employed were obviously higher for the short-term recipients than for the long-term recipients. This is related to the degree of welfare dependency of short-term and long-term recipients. Long-term recipients are more reliant on welfare assistance. Compared with the AFDC group, the FTP program also had a larger impact on the

[13] I omit the irregularly part-time employment type since for both FTP and AFDC groups the probabilities can be hardly distinguished from each other based on different characteristics. The relevant statistics are available upon request.

probability of part-time employment for the long-term recipients up to quarter 11, while for the short-term recipients the FTP program had a larger impact on full-time working rather than part-time employment. At the end of the follow-up period, the probabilities of being full-time or part-time employed were quite close for short-term and long-term recipients. It corresponds to the previous conclusion that at the end of the follow-up period the program gains became less clear as the AFDC group caught up due to higher exposure by widely publicized information of changes in state and national welfare policies and the overall healthy economic conditions.

Figure 1.8: Employment Probability for Long-Term Recipients

(After Random Assignment)

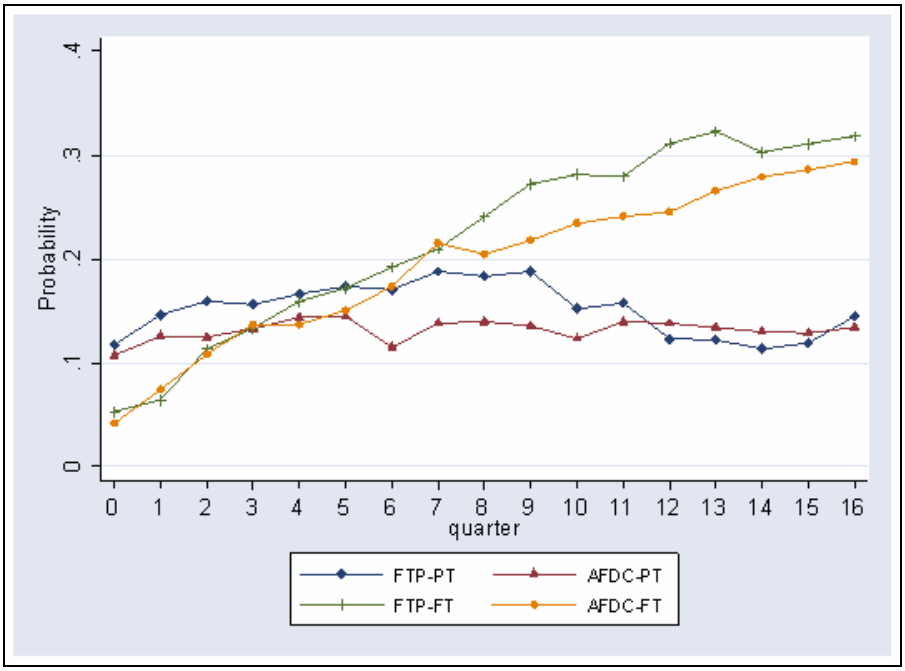


Figure 1.9 and Figure 1.10 plot the probabilities of employment for the first time recipients and non-first time recipients. At the time of the random assignment, the first time recipients were more likely to work part-time and full-time. One puzzling result from Figure 1.9 is that the probability of full-time employment for the first time recipients in the FTP group was clearly much lower than those for their counterparts in the AFDC group. One reasonable explanation might be

attributed to the different social enhanced services provided by FTP and AFDC program. For first-time recipients, enhanced traditional social services provided by the AFDC program and the low-income eligibility threshold were not as attractive as finding a full-time job. For those in the FTP program, the enhanced social services program including pre-job training, employment related services, intensive case management, etc., encouraged them to take advantage of the FTP program in order to accumulate more human capital for future labour market participation. Compared with the non-first time recipients, the first time recipients were apparently more likely to be full-time employed and less likely to take the part-time job, no matter which group they were in.

Figure 1.9: Employment Probability for First-Time Recipients (After Random Assignment)

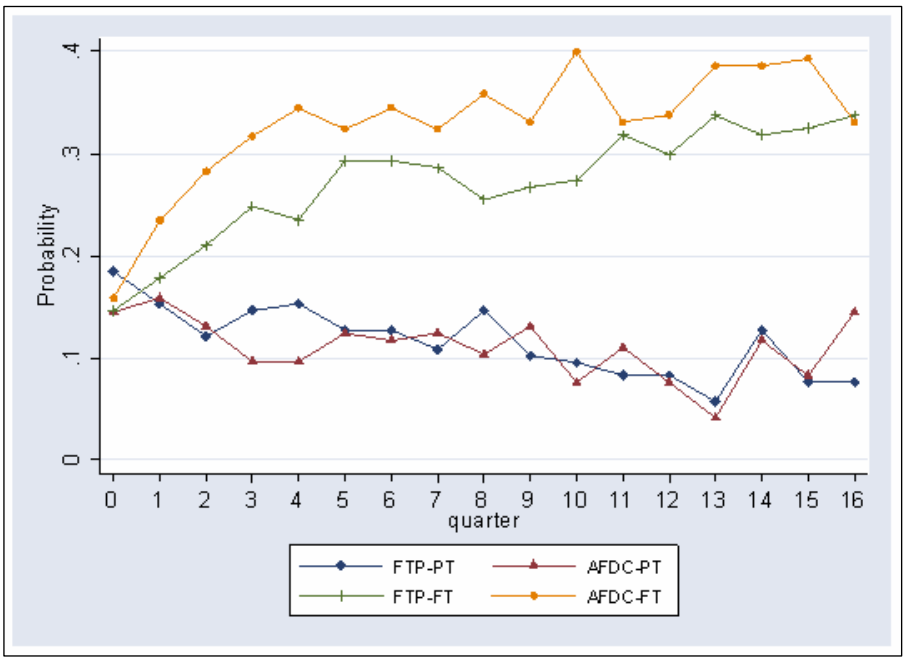


Figure 1.10: Employment Probability for Non First-Time Recipients (After Random Assignment)

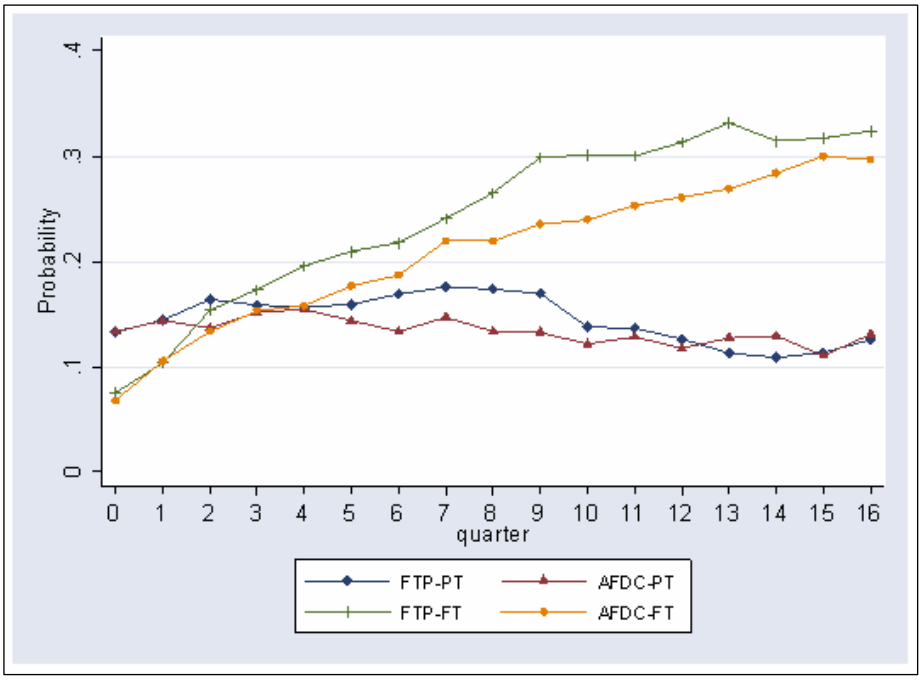


Figure 1.11 and Figure 1.12 plot the probabilities of employment for the families with and without pre-school age children[14]. There are 1807 families with preschool children, more than half of the total observations in the data. For the families without pre-school age children in either FTP or AFDC groups, the probabilities of full-time employment were higher than those with pre-school age children. Both graphs show that the FTP program had some impact on the probabilities full-time employment, especially from around year 2 up to the end of the study period. For example, the program can account for about 5% of the increase in the full-time employment probability for those with pre-school age children in quarter 9 and 13, and about 9% of the increase for those without pre-school age children. The impact was slightly larger for those without pre-school age children during the next half of the study period. Considering the more generous transitional child-care assistance provided by the FTP program- the families leaving

[14] Pre-school age is defined as age 6.

Figure 1.11: Employment Probability for Families with Pre-School Age Children (After Random Assignment)

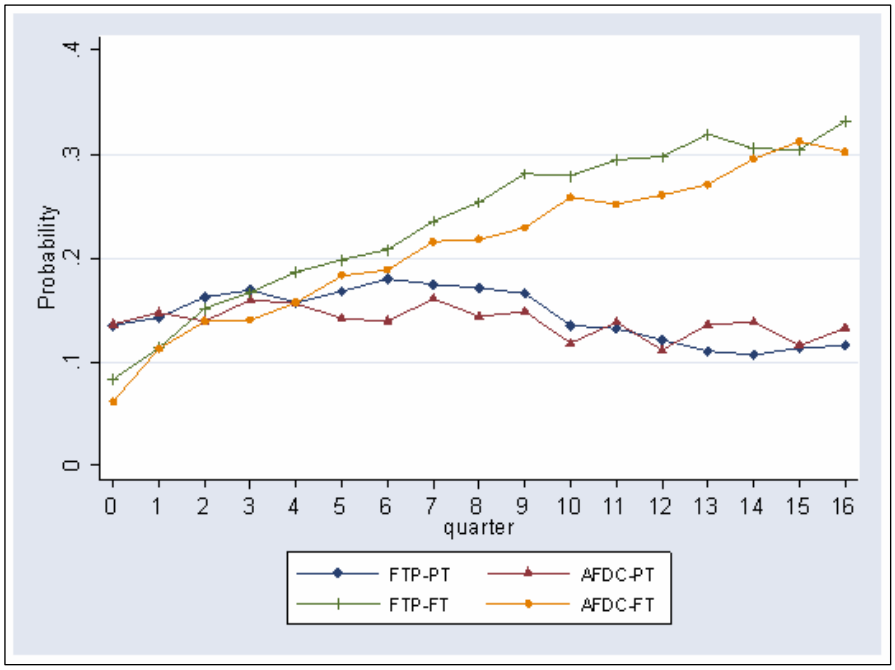
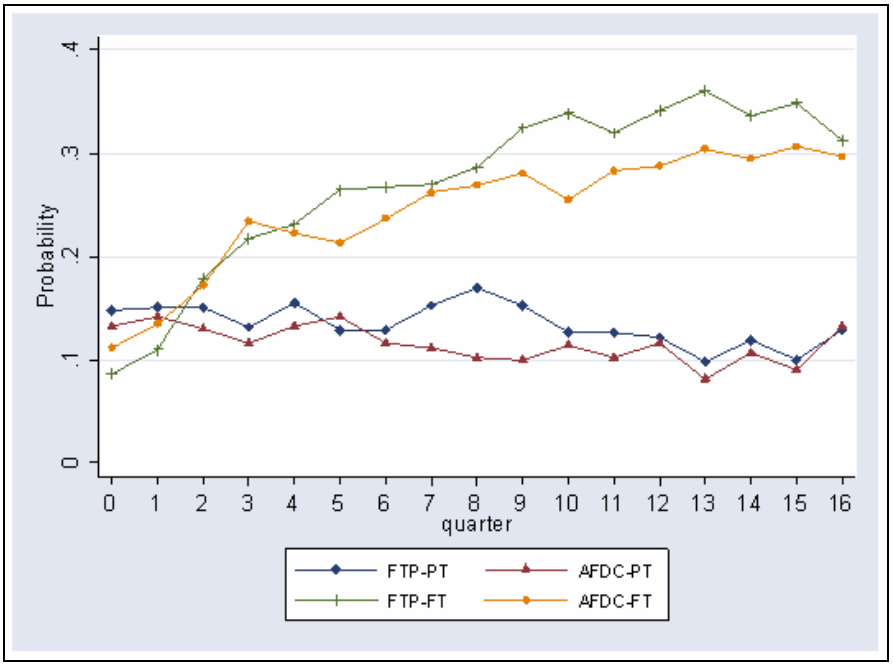


Figure 1.12: Employment Probability for Families with no Pre-School Age Children (After Random Assignment)



welfare for work in the FTP group would have 2 years of transitional child-care assistance while for those in AFDC group only one year of assistance available- it is surprising to see that the program was less effective among families with pre-school age children than among families without pre-school age children, provided that the effects of time limits, financial incentives and other work-requirements were the same for these two subgroups.

The last characteristics I examine are subgroups of high and low levels of economically disadvantaged families[15]. By definition, the high level group had higher degrees of welfare dependency, faced more serious job barrier, and would have lower earning potential than the low level group. Figure 1.13 and Figure 1.14 plot the probability of employment for high-level and low-level of economically disadvantaged families respectively. It is not surprising to see that at quarter 0, or the time of random assignment, the probabilities of employment were quite low and nearly none of the high level of economically disadvantaged individuals was employed full-time. While the program impact became apparent for the low level group in terms of full-time employment, it is difficult to be distinguished among the high level group. Thus the employment gain was more concentrated among the low level of economically disadvantaged families. The probabilities of full-time employed were increasing from 25 percent for both FTP and AFDC groups to 48 percent for the FTP group and to 45 percent for the AFDC group. The program failed to improve the welfare of the high level group and this subgroup should deserve more attention due to their relatively bigger disadvantages in the labour market.

[15] High-level is defined as level 1 and Low-level as level 3 in Table 3.

Figure 1.13: Employment Probability for H-Level Economically Disadvantaged (After Random Assignment)

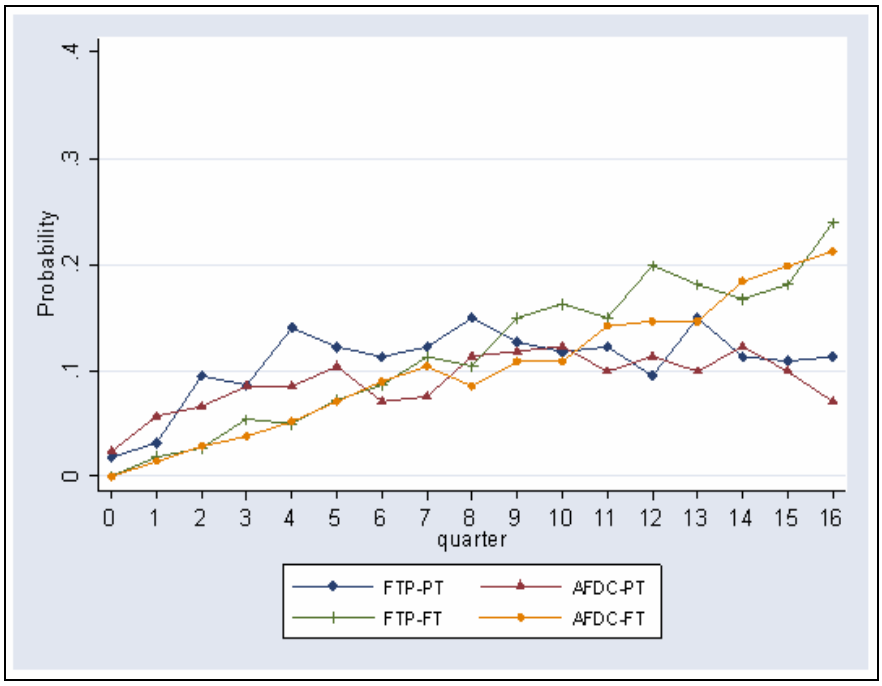
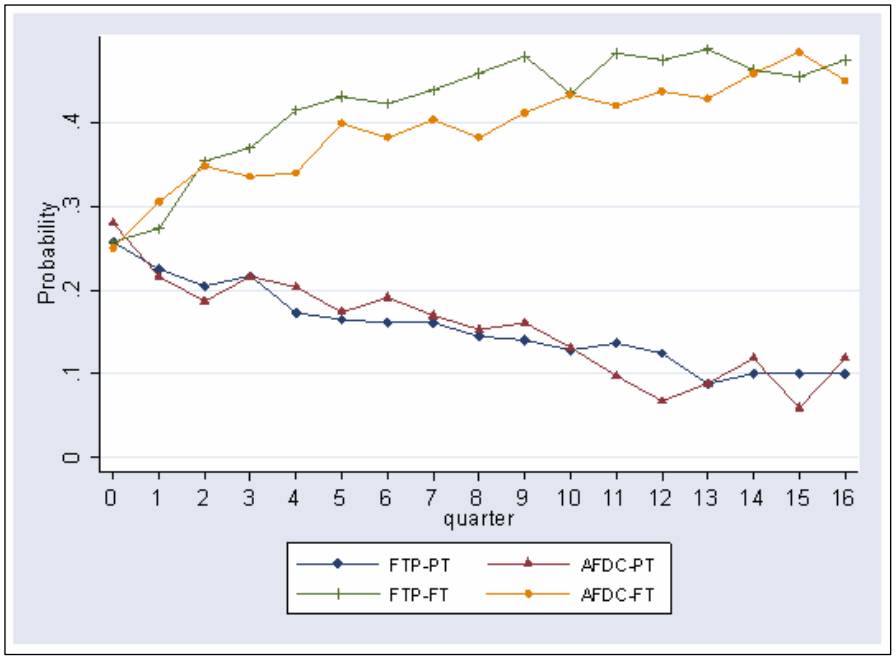


Figure 1.14: Employment Probability for L-Level Economically Disadvantaged (After Random Assignment)



From the previous analysis that was based either on the whole FTP and AFDC group or the subgroups of different characteristics, one common characteristic is that the FTP program had impact mostly in year 2 and 3, but the program gain in terms of employment became quite small and even indiscernible at the end of the study period. Over the 16 quarters, the single mothers were more likely to switch from part-time jobs to full-time jobs. The employment gain from the implementation of the FTP program was mainly in full-time employment. However, for some subgroups, I have found that the transitional child care assistance provided by the FTP program didn't help single mothers with pre-school age children to be better off in the job market, and the program showed poor outcomes on improving the subgroup of highly economically disadvantaged members. Since 4-year followed up period is too short to make credible conclusions about the long-term employment dynamics, I am only able to infer that at least in the short term, the FTP program as a whole had some positive effect on increasing full-time employment probabilities, but this effect diminished at the end of the study period.

IV. Unemployment Spells and Econometric Estimation

A. Econometric Model

The previous section is mainly devoted to analyzing the probability of three different types of employment within 16 quarters after the random assignment. A further step is to look at how the FTP program has changed the speed of leaving unemployment or consequently picking up a new job that is irregularly part-time, part-time, or full-time one. Here I refer to the hazard rate (or transition rate) of exit from unemployment, which is a conditional probability that describes the incidence of an event over time. It differs naturally from the previous probability analysis that simply takes means of incidence rates for each time interval without considering the prior conditions.

I estimate a discrete-time proportional hazard model with competing risks. As the sample members were mostly characterized by low education and weak labour force attachment, their job stability was relatively low. Therefore they

frequently switched between working and not working. I could obtain at least one spell for each individual during the study period. Nevertheless, since the treatment dummy, or the FTP program, is of our interest for the moment, I only take the single unemployment spell that ends at or after the random assignment. Due to this nature, the model actually estimates the program impact on the probability of taking up the first available job after the random assignment.

Let T denote the nonnegative length of the unemployment spell as I construct in the data. x(t) is a vector of covariates summarizing observed differences between individuals at time t. x(t) may be time varying. Let J denote the competing risks or destination states which in the data is irregularly part-time employment / part-time employment / full-time employment. The discrete time hazard (or transition) rate at tth quarter from unemployment to another destination is the conditional probability of the exit from unemployment at t, given that she has been unemployed until (t-1)th quarter and is defined as:

$$h_j(t|x(t)) = \Pr(t-1 < T \le t | T > t-1, J = j, x(t)), \; j = 1,2,3$$

and $0 \le h_j(t|x(t)) \le 1$.

A discrete time representation of a continuous time proportional hazards model leads to the so-called complementary log-log specification[16]. The resulting discrete time hazard rate is specified as

$$h_j(t|x(t)) = 1 - \exp[-\exp(\gamma_{jt} + x(t)'\beta_j)]$$

or

$$\log(-\log[1 - h_j(t|x(t))]) = x(t)'\beta_j + \gamma_{jt}$$

where $\gamma_{jt} = \log\left[\int_{t-1}^{t} h_{j0}(u)du\right]$ is the log difference between the integrated baseline hazard $h_{j0}(t)$ evaluated at quarterly interval (t-1, t]. To put things another way, the γ_{jt} summarizes the pattern of duration dependence in the interval, here quarterly, hazard rates. It can be estimated by a parametric function of duration or

[16] see Prentice and Gloeckler(1978)

non-parametric specification[17]. β_j represents a vector of the destination specific parameters of the covariates.

Up to now, the econometric model discussed assumes no unobserved individual effects. Let *v* denote the unobserved differences between observations. The random variable *v* takes a probability distribution with mean 1 and finite variance σ^2. The crucial assumption about *v* is that it is distributed independently of *X*, the covariates vector.

For the discrete time proportional hazard model, the Gamma distribution has been the most popular distribution used in the literature. The hazard function incorporating unobserved heterogeneity, *v*, is written as:

$$h_j(t|x(t),v) = 1 - \exp\left[-\exp(\gamma_{jt} + x(t)'\beta_j) + \ln(v)\right]$$

or

$$\log(-\log\left[1 - h_j(t|x(t))\right]) = x(t)'\beta_j + \gamma_{jt} + u$$

with $u \equiv \ln(v)$, which is random variable with a mean of zero. The random variable *v*, or equivalently *u*, may be interpreted in several ways. The most common one is that it summarizes the impact of omitted variables on the hazard rate. This problem arises either because the missing variables are intrinsically unobservable or simply unobserved in the data set to hand. An alternative interpretation can be offered in terms of error of measurement in recorded variables or recorded survival times[18]. The literature suggests several findings if the unobserved heterogeneity is ignored in modelling. First of all, the 'no-frailty' model will over-estimate the degree of negative duration dependence in the hazard and under-estimate the degree of positive duration dependence. Second, the proportionate response of the hazard rate to a change in a covariate is under-estimated[19].

---

[17] A non-parametric approach to characterize the frailty distribution was pioneered in the econometric literature by Heckman and Singer (1984). The idea is essentially that one fits an arbitrary distribution using a set of parameters. These parameters comprise a set of 'mass points' and the probabilities of a person being located at each mass point.

[18] See Lancaster 1990, Chapter 4.

[19] In the statistics software STATA, pgmhaz8 provides a likelihood ratio test for checking whether the unobserved heterogeneity is statistically significant.

In the hazard model with competing risks, provided that the risks are independent, the estimation of can be simply done by estimating a single risk model. The discrete hazard rate for exit at time *t* to any destination is the sum of the destination-specific discrete hazard rates. That is:

$h(t) = h_1(t) + h_2(t) + h_3(t)$,

The likelihood function is the product of the likelihood to each of the three destinations and the likelihood of right censored, that is, still remained unemployed at time *t*.

B. Descriptive statistics of the unemployment spells.

This subsection describes some characteristics of the unemployment spells. Table 1.4 provides a comparison of the basic features of the unemployment spells and the exit destinations between the FTP and AFDC groups. There are total 2657 spells in the dataset, 1336 or FTP group and 1321 for the AFDC group. The mean spell length is 8.9 and 9.4 quarters for the FTP and AFDC group, respectively. The longest spell length is 25 quarters, which is the whole study period. These are so called right-censored spells as individuals still remained unemployed at the end of the study period. There are 313 right-censored spells out of a total of 2657. A first glance at Table 1.4 does not show obvious differences for the frequencies of exit into irregularly PT job, PT job, and FT job between the FTP and AFDC group. For exit into irregularly PT jobs, both groups were very close, with 58.9 percent for the FTP and 58.2 percent for the AFDC group, which means quite a large proportion of the members worked irregularly once unemployed and earned a very small amount. For exit into a PT job, the gap was somehow bigger, with 20.4 percent for the FTP and 18.7 percent for the AFC group. For the exit into a full-time job, the FTP group had even lower frequency than the AFDC group. However, the above information is not quite contrary to Figure 1.3 through Figure 1.5 in the previous sections. Since the spells I analyse here are the length of unemployment up to the first job after the random

assignment, they do not reflect the dynamics of switching between unemployment and employment over the whole study period.

Table 1.4 -- Comparison of Spell Features of The FTP and AFDC Groups

Features	FTP Group	AFDC Group	Total
Number of spells	1336	1321	2657
Mean spell length (in quarters)	8.9	9.4	9.2
Right-censored spells	14.2	15.5	14.9
Unemployed throughout the whole study period	11.5	12.1	11.8
Exit into irregularly PT job	58.9	58.2	58.6
Exit into PT job	20.4	18.7	19.6
Exit into FT job	6.4	7.5	7.0

Note: entries are frequencies except for number of spells and mean spell length.

C. Non-parametric Analysis

First I examine the speed of leaving unemployment, regardless of which exit destination. Figure 1.15 presents life-table estimates of the survival probability of being unemployed for the FTP and AFDC group[20]. For both groups, it shows that the probabilities declined more rapidly in the first 20 quarters. Around 75 percent members in both groups remained unemployed at the very beginning of the unemployment phase. After 20 quarters unemployed, around 18 percent of the AFDC members survived, i.e., were still unemployed and around 15 percent of the FTP members were unemployed. Though the survival probabilities for the FTP group were consistently lower than those for the AFDC group, the difference between them was not very distinct. Up to the 8th quarter, the survival probabilities for two groups traced each other quite closely. It means that for single mothers who have been unemployed less than 2 years, the rates of exit from unemployment were quite close, regardless of which group they were

[20] J. Wolff (2002): "ltable1: STATA module to overlay Survival/Hazard-curves of distinct sub-samples".

assigned. There is a kink at the 8th quarter of unemployment, with the probabilities of remaining unemployed decreasing at a faster speed and the speed of exit from unemployment starting lightly accelerated with the introduction of the FTP program. This same kinked pattern also shows up in the hazard rate plots. For those who have been unemployed recently for more than 20 quarters, the speed of exit from unemployment slowed down. The probabilities of remaining unemployed were 14 percent for the FTP group and 15 percent for the AFDC group. Overall, comparing the survival probabilities of unemployment between the FTP and AFDC group shows that the FTP program didn't significantly speed up the single mothers to take the first available job after the random assignment. The likelihood ratio test statistics of homogeneity between these two groups yields a p-value of 9 percent.

Figure 1.15: Survival Probabilities of Unemployment

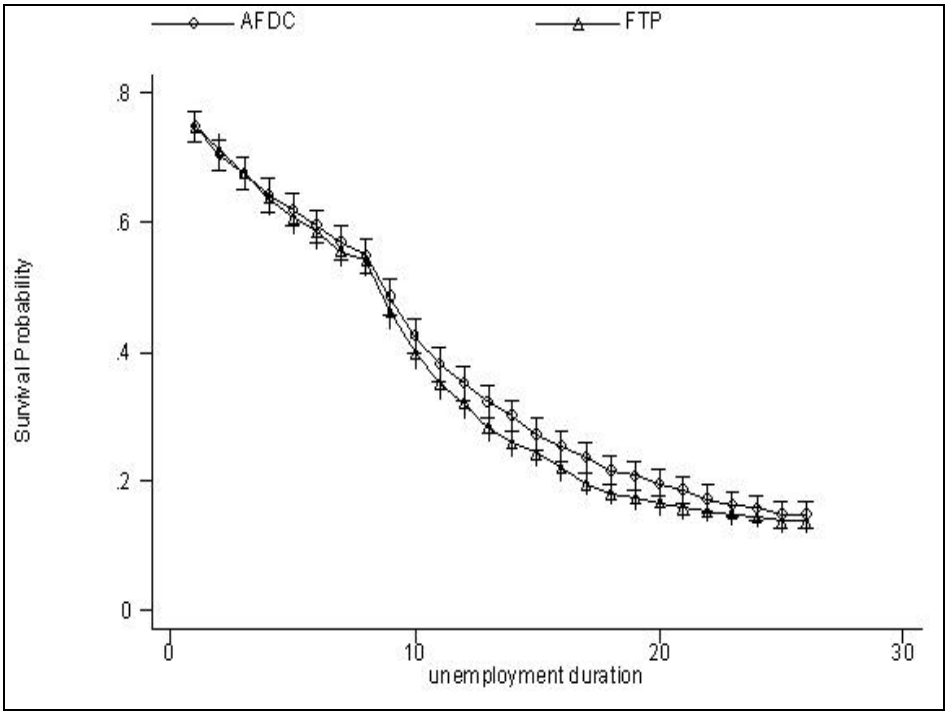


Figure 1.16 to Figure 1.18 plot the life-table estimates of hazard rates into three destinations together with 95 percent confidence intervals. A first glance at

the confidence bands show that apparently there is no significant difference between the exit rates to all three destinations for the FTP and AFDC group.

In Figure 1.16, the hazard rate of transition into irregularly part-time jobs was relative high at the first quarter of unemployment- around 28 percent for both groups. Then there is a sharp drop in the second quarter to around 4 percent for the AFDC group and 3 percent for the FTP group. As I defined irregularly part-time employment as having quarterly earning less than $510, a rather low amount, being irregularly employed in this case might mean only getting some hourly job from time to time and there might be some time spent on job search as well. Therefore, the first quarter of unemployment serves as a transition period to getting a part-time job or full-time job. Another jump of the hazard rate of exit into irregularly part-time job happened after an unemployment duration of 8 quarters, from 1 percent to 11 percent for the FTP group and from 1 percent to 14 percent for the AFDC group. It means that after two years of unemployment the probability of picking up some irregularly part-time job became higher. Since the 10^{th} quarter of unemployment, the probabilities of exit into an irregularly part-time job became less than 6 percent and remain quite constant over the rest of the unemployment horizon.

Figure 1.16: Hazard Rate of Exit into Irregularly PT Job

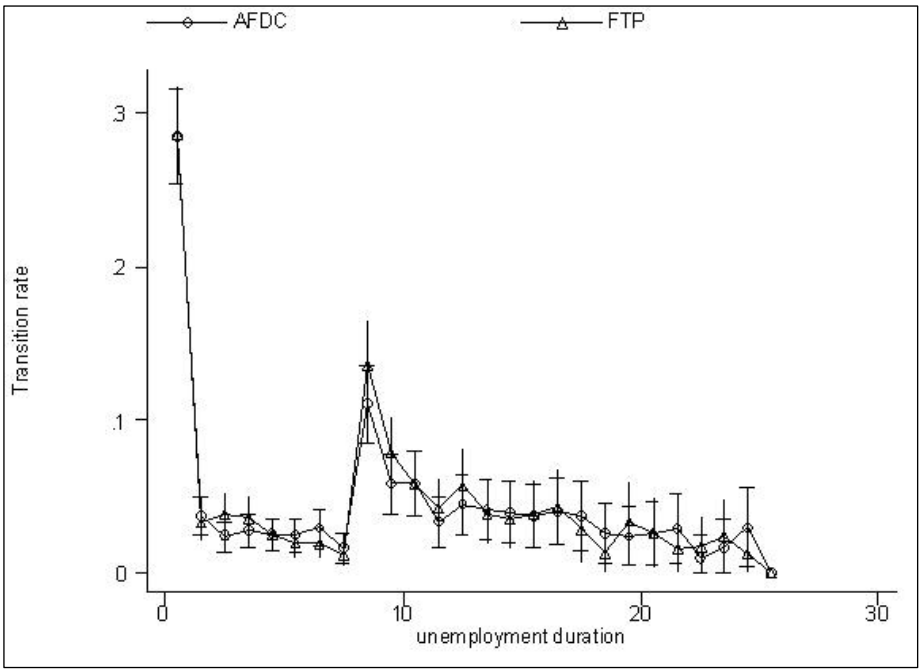
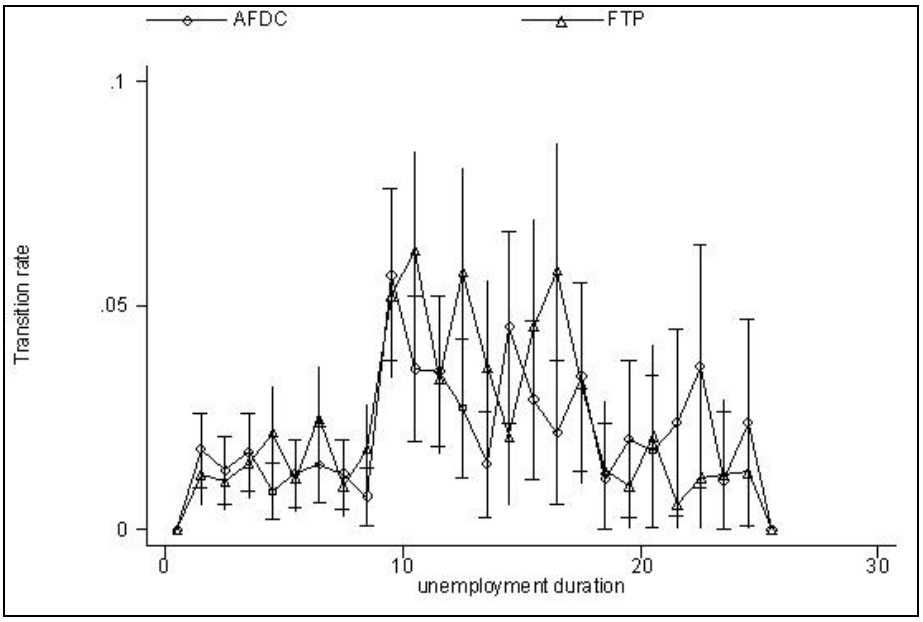


Figure 1.17 plots the hazard rates of exit into a part-time job for the FTP and AFDC group. Over the whole unemployment duration, the hazard rates, or conditional probabilities were quite low and the difference between these two groups is hardly seen. Up to the 9th quarter of being unemployed, the hazard rates varied between 1 to 3 percent. In quarter 10, the hazard rates of exit into part-time jobs increased for both groups to above 5 percent. After that the gaps between these two groups became larger. However the overlapping confidence bands show that the difference is insignificant. At the few quarters close to the end of the duration I studied, the hazard rates were a little bit higher for the AFDC group.

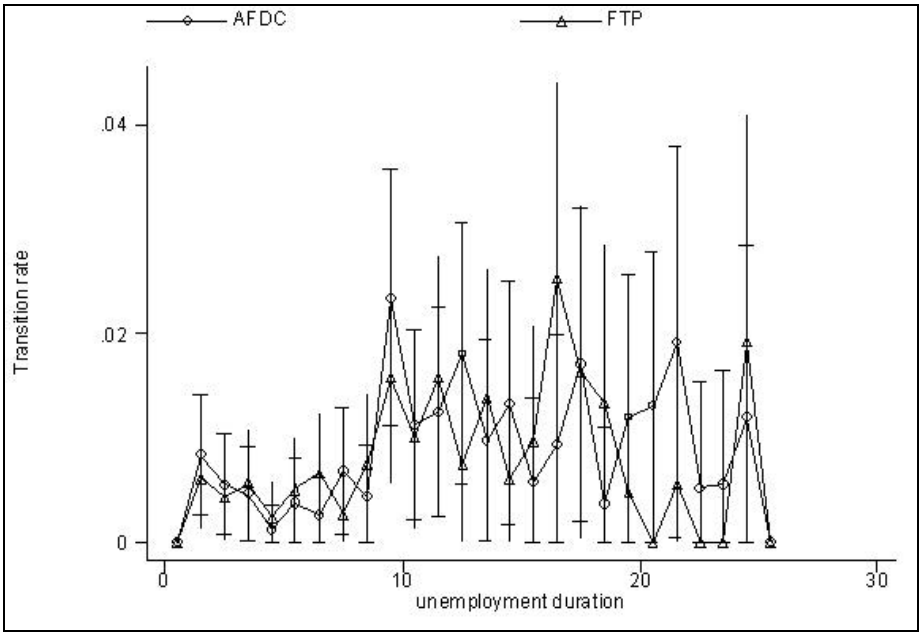
Figure 1.17: Hazard Rate of Exit into PT Job



For the hazard rates of exit into a full-time job, the pattern exhibited in Figure 1.18 is quite similar to the one in Figure 1.17. Up to the 10th quarter, the hazard rates stayed under 1 percent. In quarter 10, the hazard rates increased to 2.2 percent for the AFDC group and 1.7 percent for the FTP group. Similar to the figure for exit into a part-time job, since the 19th quarter the hazard rates for the AFDC group were higher than the FTP group except in quarter 25.

In general, the FTP didn't seem to speed up the take-up rates of the first available job compared with the AFDC group as shown in these graphs from life-table estimates.

Figure 1.18: Hazard Rate of Exit into FT Job



D. Parametric analysis of unemployment duration with covariates

In this section I discuss the effect of covariates on the hazard rates in a discrete time proportional hazard model incorporating unobserved heterogeneity that has a gamma distribution. The previous non-parametric analysis shows that there is no obvious gain from the FTP program. The estimation of the econometric model presented in subsection A controls some observed differences such as race, education, mother's age, number of children at home, whether there are pre-school age children present, welfare dependency, length of employment in the previous two years, level of economically disadvantaged, etc. The baseline hazard is parameterized by taking the logarithm of the unemployment duration and it summarizes the duration dependence.

In the study of unemployment duration, the receipt of unemployment insurance is usually an important explanatory variable. In our dataset, as the

subjects are low educated single mothers with weak labor force attachment, the unemployment insurance receipt is not a common practice. During the whole study period, 2 years prior to and 4 years after the random assignment, there are only around 1 percent of the sample members receiving unemployment insurance. Moreover, there is no information for calculating the remaining time eligible for unemployment insurance receipt. I leave this variable out of the estimation.

One shortcoming of our data is that the characteristics of the exit states are missing. Therefore I have no detailed information such as the industry, actual hourly pay, whether health insurance is provided, distance from home, etc, about the job they took. From the previous analysis, I would predict that the jobs they took are mostly low paying and less stable ones in line with their background.

Table 1.5 displays the coefficient estimates from the model with unobserved heterogeneity together with the log likelihood and the likelihood ratio test of heterogeneity. The coefficient of log(t) is 1.96 for exit into irregularly part-time jobs, which can be interpreted as the elasticity of the hazard with respect to time. A similar interpretation applies to baseline hazards of the exit into part-time (2.15) and full-time jobs (2.41), respectively. As this negative duration dependence pattern shows, the hazard rates are getting higher with time. After controlling for the observed characteristics and unobserved heterogeneity, the coefficient on the FTP dummy is insignificant for exit into all three destinations, which is in line with the analysis from previous graphs. The FTP program in general didn't speed up the exit from unemployment to irregularly part-time jobs, or part-time jobs or full-time jobs. The same true for the insignificant impact of 36-month time clock[21]. The race dummy of black appears only significant in the exit into irregularly part-time jobs. The coefficient, 0.69, means that the hazard for black single mothers is on average 31 percent lower than their white counterparts. For exit into part-time or full-time jobs, there is no significant difference between being black or white. Number of children and whether having pre-school age children at home are only significant in exit into part-time jobs. The single

[21] This is actually an imputed variable as the AFDC members faced no time limit at all. Michalopoulos from MDRC provided detailed explanation about imputation.

mothers preferred part-time jobs when there were more kids at home. For mothers who were under 35 years old at the time of random assignment, the hazard rates are significantly higher for exit into a non full-time job than those who were older than 35 years old. Previous employment in the 2 years preceding the random assignment is important for increasing the hazard to exit from unemployment into three different jobs. The indicator of less economically disadvantaged is significant in explaining the hazard of exit as well. The likelihood ratio tests yield chi-squares that are significant at 1 percent level, which indicate the unobserved heterogeneity can't be ignored in my model.

Overall, the estimation from the frailty model indicates that the FTP program had no significant impact on exit from unemployment to employment in the first available job after the random assignment. As I study only the single unemployment spells during which the random experiment took place, this insignificant result is not completely surprising. For the single mothers assigned to the FTP program, the incentive to take up a job immediately after the random assignment may not be strong since they might prefer to stay unemployed on welfare and take advantage of the job-related services furnished by the FTP program to accumulate the human capital for future use. Therefore, I could not capture the potential impact of the program in the long run due to the data limitations.

V. Conclusion

In this chapter, I mainly focus on how the FTP program influences the probabilities of employment for single mothers, who are the major welfare recipients, and the speed of leaving unemployment to start the first available job. The three main components of the FTP program include time limits, financial incentives and work-related mandatory requirements. I study the effect of the program as a whole instead of a specific policy component in this chapter.

I found that over the 4-year follow-up period, the FTP members were more likely to be full-time employed and less likely to be part-time employed. There is no significant difference in terms of irregularly part-time employment,

which is possibly due to the nature of low quarterly earning of this kind of employment. At the end of the follow-up period, the employment gain started diminishing and these two groups became closer in employment probabilities. Other interesting results I find include that the more generous transitional child care assistance provided by the FTP program did not increase the employment probabilities of the single mothers with pre-school age children relative to those without, and that the FTP program was not effective in improving the economic situation of those highly economically disadvantaged members. Though the short study period does not allow me to make long term conclusions such as the impact on the employment in 10 years, it is no doubt quite helpful in understanding the impact of the FTP program in its implementation period from 1994 to 1999.

The discrete time proportional hazard model with unobserved heterogeneity yields insignificant estimates of the effect of the FTP program on the hazard rates of exit from unemployment into irregularly part-time jobs, part-time jobs or full-time jobs. The probability difference of remaining unemployed after certain time elapses between the FTP and AFDC groups are statistically insignificant. Once assigned to the FTP group, the single mothers tended to stay on welfare at the onset and took advantage of the enhanced services since their low education, weak labor market attachment and child care problems presented a serious job finding barrier. This partly helps explain the insignificant impact of the FTP program on the exit from unemployment to the first available job. Even though the dynamics of switching between unemployment and employment is taken into account, the FTP program is expected to exert no important influence due to the limited potential job offers faced by the single mothers.

Table 1.5: Frailty Model Estimates of Program Effect on Hazard Rates for Exit into Three Destinations

Variable	Exit into Irregularly PT Job (1)	Exit into PT Job (2)	Exit into FT Job (3)
Constant	-7.877** (0. 924)	-10.948*** (0.918)	-13.2705*** (1.676)
Baseline hazard (log(t))	1.960*** (0.211)	2.146*** (0.242)	2.411*** (0.420)
FTP dummy	0.184 (0.179)	0.227 (0.158)	-0.218 (0.281)
36-month time limit	0.234 (0.222)	-0.024 (0.196)	0.177 (0.347)
Mother black	0.692*** (0.200)	-.020 (0.172)	-0.041 (0.299)
Number of children	-0.059 (0.089)	0.195*** (0.073)	0.158 (0.138)
Year of schooling	-0.140** (0.070)	0.077 (0.057)	0.144 (0.100)
Having pre-school age children	0.060 (0.232)	-0.496** (0.208)	-0.355 (0.372)
Mother age less than 20	1.926*** (0.644)	1.588 *** (0.567)	1.123 (1.028)
Mother age between 20 and 24	1.845*** (0.559)	1.258 *** (0.480)	0.695 (0.864)
Mother age between 25 and 34	1.219** (0.524)	0.752* (0.444)	1.060 (0.796)
Mother age between 35 and 44	0.391 (0.535)	-0.222 (0.460)	0.844 (0.805)
Less economically disadvantaged	6.773*** (0.734)	1.310*** (0.472)	2.673*** (0.938)
Short-term recipient	-0.966*** (0.251)	-0.395 (0.213)	-0.028 (0.356)
Month of welfare use in 24 months prior to random assignment	0.007 (0.011)	.030*** (0.010)	0.004 (0.016)
Quarters of employment in 2 years preceding random assignment	2.246*** (0.232)	0.738*** (0.074)	0.745*** (0.124)
Log likelihood	*-4274.7838*	*-2322.7533*	*-1011.4948*
LR test of Gamma var. = 0, chibar2(01)	*688.24*	*75.77*	*39.5479*

Note: *** 1 percent significance level, ** 5 percent significant level, * 10 percent significance level.

Reference:

Blank, Rebecca M (2000). "*What Causes Public Assistance Caseloads to Grow?"* Journal of Human Resources, 2001, Vol.36(1,Winter), Pp.85-118.

Bloom, Dan et all (2000). *"The Family Transition Program: Final Report On Florida's Initial Time-Limited Welfare Program",* MDRC

Council of Economic Advisers (1997). *Technical Report: Explaining the Decline in Welfare Receipt,* 1993-1996. Washington, DC: Council of Economic Advisers.

Cox, D.R (1972). *"Regression models and life-tables"*, Journal of Royal Statistical Society B., 34, 187-202

Grogger, Jeff and Michalopoulos, Charles (2003). *"Welfare Dynamics under Time Limits",* Journal of Political Economy, Vo.111, Issue 3, Pp.530-554.

Jenkins, S.P (1997): "*Discrete Time Proportional Hazards Regression (pmghaz)"*, Stata Technical Bulletin, 39, 22-32

Jenkins, S.P (2004): *Survival Analysis,* Manual

Kemple, James J. and Joshua Haimson. *Florida's Project Independence: Program Implementation, Participation Patterns, and First-Year Impacts.* New York: Manpower Demonstration Research Corporation, January 1994.

Lancaster, T. (1990). *"The Econometrics Analysis Of Transition Data*." Cambridge University Press, Cambridge.

Martini, Alberto and Michael Wiseman (1997). *"Explaining the Recent Decline in Welfare Caseloads: Is the Council of Economic Advisers Right?"* Washington, DC: The Urban Institute, July 1997.

Meyer, B.D (1990). "*Unemployment Insurance And Unemployment Spells*". Econometrica, 58(4), 757-782

Moffitt, Robert (1983). "*An Economic Model of Welfare Stigma*." American Economic Review Vol. 73, 1023-1035.

Moffitt, Robert. *"Incentive Effects of the U.S. Welfare System: A Review*." Journal of Economic Literature. Vol.30, March 1992, 1-61.

Moffitt, Robert. *"Time Limits."* Mimeo, Johns Hopkins University, November 1998.

O'Neill, June, Lauri Bassi, and Douglas A Wolf. *"The Duration of Welfare Spells.*" Review of Economics and Statistics. Vol. 69, May 1987, 241-248.

Swann, Christopher A. *"Welfare Reform When Agents are Forward-Looking."* Mimeo, University of Virginia, December 1998.

Ziliak, J.P., Figlio, D.N., Davis, E.E., and Connolly, L.S. "*Accounting for the Decline in AFDC Caseloads: Welfare Reform or Economic Growth?",* Journal of Human Resources, 35(3), 570-586, Summer 2000

Appendix 1. Florida's Family Transition Program:
Comparison of Policy Components Between FTP and AFDC

Characteristics	FTP Group	AFDC Group
Time limits on cash assistance receipt	24 months in any 60-month period for most recipients; 36 months in any 72-month period for the least job-ready. Exceptions under certain circumstances.	None.
Financial work incentives	The first $200 earnings are disregarded. The benefit reduction rate, or tax rate, for the remaining earnings is 50 percent.	First 4 months: $120 disregard and 67 percent benefit reduction rate. Month 5-12: $120 disregard and 100 percent benefit reduction rate. After month 12: $90 disregard.
Mandatory work-related activities (MWRAs)	30 hours per week of either work or word related activities. Mandate could be satisfied by participating in a welfare-to-work program that provided enhanced employment and training services. Exemptions only for mothers with infants under 6 months of age.	30 hours per week of either work or work-related activities. Mandate could be satisfied by participating in a welfare-to-work program that provided conventional AFDC services. Exemptions for mothers with children under 3 years old.
Child care assistance for families leaving welfare for work	Two years of transitional child care assistance; eligibility beyond that point depended on eligibility for other programs.	One year of transitional child care assistance; eligibility beyond that point depended on eligibility for other programs.
Parental responsibilities mandates	Parents had to ensure that children attended school regularly, and had to speak with teachers at least once each grading period. Applicants with preschool children had to prove that children had begun immunizations.	None.
Asset limit for cash assistance eligibility	$5,000	$1,000

Source: The Family Transition Program: Final report on Florida's Initial Time-Limited Welfare Program, summary report, sum-3

Chapter Two

Accounting for Life Insurance Holdings: Evidence from German Socio-Economic Panel Studies

Life insurance is used to insure against the risk of premature death in order to smooth the consumption. This chapter studies the demand for life insurance holdings in Germany within an economic framework. I examine the driving forces behind the life insurance demand with the data from German Socio-Economic Panel Studies. After correcting for the autocorrelation of disturbances in the unbalanced panel data, I find that the bequest motives are quite strong among the households with children in holding life insurances; being unemployed has significant negative effect; higher marginal tax rates are related to higher demand for life insurances due to the tax incentive scheme in Germany. The self-reported risk attitude captures the individual heterogeneity in the panel dataset. I also evaluate the recent abolishment of tax advantages on life insurance contracts signed after Year 2004 and find positive anticipation effects in year 2003 and year 2004 after the beginning of the policy discussion.

I. Introduction

Economists frequently refer to consumption smoothing, which means the household tries to maintain the same living standard level in the presence of lifetime uncertainties. To smooth consumption, the economic theory predicts that the household (with the key rationality assumption) will save and insure in order to reduce the risk of variability in their living styles. Life insurance, therefore, like other insurance, provides protection against the consumption risk in the event of a premature death of a household member. The first classical paper, written by Yaari (1965), describes the relationship between life insurance and uncertainties, where the bequest motive is modelled as one of the driving forces behind buying life insurance. Furthermore, a set of conditions are derived under which the household will be fully insured against the lifetime uncertainties.

Life insurance is an important means of saving in Germany. All life insurance policies provide term insurance, which refers simply to insurance provided by the policy in a given year, and pay death benefits to surviving family members or other beneficiaries. Whole life insurance combines term insurance with a savings plan and can generate annuities or lump sum payments after a certain amount of time (usually 12 years). Paying premiums for a whole life policy is therefore equivalent to contributing to one's saving account and also buying annual term insurance.

The significance of life insurance as part of security for old age, disability and surviving dependents has been increasing for many years. According to GDV (Gesamtverband der Deutschen Versicherungswirschaft[22]), in Year 2003 there were around 8.0 million new contracts concluded. The premium income of life insurance increased from 13.2 billion Euro in Year 1980 to 67.3 billion Euro in Year 2003 and it amounts to more than 45% in private household savings. The share of life insurance in total volume of pension provision increases from 19.0% in Year 1993 to 33.3% in Year 2003.

[22] German Insurance Association

Why do Germans spend so much on life insurance? What are the driving forces behind this demand? I consider the following scenarios: changes in family structure, changes in labour market participation and changes in the tax code.

Changes in family structure, specifically changes in marital status and the presence of children at home can change bequest motives. A bequest motive is often viewed as a 'joy of giving'. It was first seen in Yaari (1965). Bequest motives can be linked to the altruism of the insured towards their surviving spouse or their descendents. Several papers also depict how life insurance holdings vary across different household types. Auerbach and Kotlikoff (1991) find that middle-aged married couples are more likely to have life insurance. Bernheim (1991) focuses on elderly married and single individuals. Chang (2004) presents a life cycle model of life insurance that takes into account the ages of young beneficiaries. They found that the size of contingent bequest shrinks as the child ages.

In terms of change in labour market participation, I consider those who lose their job as well as women who exit from the labour market and become housewives. Being unemployed imposes a big shock to the household's financial stability. When the head of the household or the spouse is unemployed, how can the household smooth its consumption and simultaneously insure life against future uncertainties? Will the household spend down its savings or cancel any insurances policies (not the obligatory ones) to secure the current financial situation? Another concern is the asymmetric gender bias within a household in terms of life insurance holdings. Are housewives less likely to be insured? If so, a husband faces a lower level of protection if his spouse passes away. Bernheim, Forni, Gokhale and Kotlikoff (2003) examine the life insurance adequacy for couples approaching retirement age with the Health and Retirement Study (HRS) data and they find that a sizable minority of couples in the HRS sample are less likely to be insured. Given the aging problem in Germany and the rest of Europe, it is important to examine the consequences of inadequate life insurance on the welfare of the elderly.

Change in tax code is positively related to life insurance holdings in Germany. As life insurance is equivalent to a tax-free saving in Germany,[23] theoretical models predict that tax incentives will increase the demand. Winter and Walliser (1998) find the demand for life insurance is positively related to the average tax rate. Richter and Ruß (2001) find that in the German insurance market, possible advantages from purchasing the contract combining whole life insurance and immediate annuity are clearly increasing in the tax rate and in age. I construct the marginal tax rate[24] out of GSOEP and use it to study the effect of tax incentives. However, tax benefits have been abolished on life-insurance policies signed after Dec. 31, 2004[25]. As the discussion was first beginning in Year 2003, I use the data to study the program anticipation effect, i.e., how people respond to their expectation of policy change.

In addition, I explore the self-reported risk preference in predicting the demand for life insurance. Dohmen et al. (2005) have proved the consistency and validity of this self-reported measure for actual behaviour. The risk preferences reported by the individuals in Year 2004 survey therefore can reflect the underlying trait and are relevant for predicting behaviour. Bonin et al. (2006) investigate whether risk preferences explain how individuals are sorted into occupations with different earning risks and they find that individuals with low willingness to take risks are more likely to be sorted into occupations with low earning risks.

I construct a longitudinal data out of German Socio-Economic Panel (SOEP). With the longitudinal data, there is sufficient variation to identify the individual effects across the time. With panel probit models correcting for auto-correlated disturbances, I find that bequest motives are quite strong among married couples with dependent children. The probability is about 14 percent

[23] German tax authorities consider an annuity payment as consisting of two portions. One portion is the pay-back of the invested amount of capital and therefore is not subject to taxation as the invested capital comes from after-tax income. The second portion is considered as interest and thus taxable.

[24] This variable was constructed out of SOEP. See Schwarze (1995) for detailed method.

[25] Payouts will be subject to personal income tax. People aged 60 and above who have invested for at least 12 years will get half the payout tax-free.

higher in the whole sample. Being unemployed significantly reduces life insurance demand by 2 percent. Housewives are less likely insured. The same is for the low-income families. The marginal tax rate is strongly related to the higher demand for life insurances. A person who is subject to the average marginal tax rate is about 5 percent more likely to buy life insurance than his counterpart who is not subject to taxation. I find evidence of the program anticipation effect as well. In Year 2003 and 2004, the demand is much higher than the year before after I control other factors. Finally, I find that the self-reported risk attitude is positively associated with life insurance holdings. The individuals who are inclined to take riskier actions are also more likely to sign life insurance contracts. It implies that the self-reported risk attitude captures the individual heterogeneity instead of a measure of risk aversion. The risk-inclined individual utilizes life insurance to minimize utility loss out of rationality.

The rest of this chapter is organized as following. Part 2 presents the theoretical model. Part 3 describes the data. Part 4 discusses reduced-form estimations and the empirical results. The last part concludes this chapter and discusses future research potential.

II. Theoretical Model

As life insurance is used to insure against the unforeseen mortality risk and to provide a way to smooth the consumption for the surviving dependents, it is seen as a way of inter-temporally allocating resources. The life cycle model is a standard way for the economists to model choices such as consumption, saving, labour supply, fertility, etc. Theoretical models of the demand for life insurance have been derived by Yaari (1965), Fischer (1973), Campbell (1980), Lewis (1989), and Bernheim (1991). The first life cycle model of life insurance was proposed by Fischer (1973) to examine the life cycle patterns of consumption, savings, and insurance purchases. Winter and Walliser (1998) derives life insurance demand in a three-period model with a 'joy-of-giving' bequest motives. In their paper, life insurance is modelled as a combination of term life insurance and a savings plan. As their model incorporates the salient features of the German

tax and pension system, I adopt their model as my basic framework and extend to allow labour market participation.

In my model, a representative individual lives *T* years and maximize the expected utility function by choosing consumption, leisure and bequest. I set the starting period to be when individual is 20 years old. T is the maximum years the individual could live. In this model, consumption, leisure, bequests are assumed to be separable in a constant relative risk aversion (CRRA) utility function.

$$U(c_t, l_t, b_{t+1}) = \sum_{t=20}^{T} \frac{1}{1-\gamma} \beta^{t-20} \left[c_t^{1-\gamma} + l_t^{1-\gamma} + \eta_{t+1} b_{t+1}^{1-\gamma} (1-\pi_{t+1}) \right] \prod_{s=20}^{t} \pi_s$$

where *c* is the consumption, *l* is the leisure, *b* represents the bequest, γ is the risk aversion parameter of the CRRA utility function, β is the discount factor, η is the weight on bequest, and π is the survival probability. Specifically, $\pi_T = 0$ as the death at the end of life course is certain.

The budget constrain before retirement can be written as:

$$c_{t|t<65} = (H - l_{t|t<65}) w_{t|t<65} (1-\tau^s) - Z_t L_{t+1|t<65} - S_{t+1|t<65} + S_{t|t<65} R^c + \alpha L_{t|t<65}$$

$$b_{t+1|t<65} = S_{t+1|t<65} R + L_{t+1|t<65}$$

where *w* is the hourly wage, H usually equals to 24 hours, τ^s is the payroll tax that is contributed to German public pension system and returns pensions in old age, L_{t+1} is the life insurance purchased, Z_t is the price of life insurance, S_{t+1} is amount invested in bonds, R^c is defined as: $R^c = 1 + r(1-\tau^c)$, with r standing for rate of return on bonds and τ^c is the capital income tax, *R* is defined as 1+*r*, α is the exogenous savings portion of the life insurance. As the life insurance in this model is whole life insurance that combines the term life insurance and a savings plan, when the policyholder survives, a fraction of the insurance sum, or the cash value, can be withdrawn.

The official retirement age is 65 and the individual receives pension after retirement. Furthermore, the insurance contracts usually allow the insured to cash out a balance up to face value of the insurance contract at retirement. Therefore, I do not add life insurance in the bequests. In addition, it is reasonable to assume

that people do not start buying life insurance when they are old as the premium would be too high to offset the benefits. The budget constraint after retirement is therefore different from that during the working period.

$$c_{t|t\geq 65} = \tau^s \sum_{t=20}^{65} (H - l_{t|t<65}) w_t G^{65-t} - S_{t+1|t\geq 65} + S_{t|t\geq 65} R^c + \alpha L_{t|t\geq 65}$$

$$b_{t+1|t\geq 65} = S_{t+1|t\geq 65} R$$

where G equals $1 + g$ and g is the internal rate of return of the pension system.

The first order conditions from the maximization problem yields the following relationships between the choice variables of consumption, leisure, and bequests for those $t<65$:

$$\frac{c_t}{c_{t-1}} = \left[\frac{1 - RZ_{t-1}}{\beta \pi_t (R^c - \alpha R)} \right]^{-\frac{1}{\gamma}}$$

$$\frac{l_t}{l_{t-1}} = \left[\frac{w_t (1 - RZ_{t-1})}{w_{t-1} \beta \pi_t (R^c - \alpha R)} \right]^{-\frac{1}{\gamma}}$$

$$\frac{c_t}{b_t} = \left[\frac{(1 - \pi_t) \eta_t (1 - RZ_{t-1})}{\beta \pi_t (R^c Z_{t-1} - \alpha)} \right]^{-\frac{1}{\gamma}}$$

$$\frac{l_t}{b_t} = \left[\frac{(1 - \pi_t) \eta_t w_t (1 - \tau^s)(1 - RZ_{t-1})}{\beta \pi_t (R^c Z_{t-1} - \alpha)} \right]^{-\frac{1}{\gamma}}$$

The maximization can be solved recursively by solving the consumption and bequest in the last period and then substituting backwards to the previous period. I will not go further into the algebraic solutions to the model but instead simply illustrate the implications of the model. First, life insurance demand depends on the weight of bequest motives η. In the empirical model, the marital status and number of children in the household can capture the bequest motives. Second, demand for life insurance is positively related to income. While labour force participation provides a source of income, and labour income uncertainties dominate the financial capital income vulnerability, the loss of labour income therefore will dampen the demand for life insurance. I examine specifically those who lost jobs and those women who exit from the labour force and become

housewives. Third, tax parameters certainly cause changes in budget constraint. Furthermore from the comparative statics, the increase in τ^c, the capital income tax rate will increase the R^c and therefore the amount of life insurance increases to maintain the smooth consumption stream in different periods. I examine the effect of the marginal tax rate on life insurance demand as well the program effect of abolishing tax benefits on life insurance. Finally, as the CRRA model implies, the risk aversion parameter γ is positively related to life insurance demand. The greater the risk aversion is, the more willing the individual to smooth consumption over life cycle, resulting in a higher demand for life insurances. In the following sections, I use the GSOEP data to test these theoretical predictions.

III. Data and Descriptive Statistics

The data source is German Socio-Economic Panel Studies (SOEP), which is equivalent to the Current Population Survey (CPS) in the U.S. The SOEP was started in 1984 as a longitudinal survey of private households and persons aged 18 years and older in the Federal Republic of Germany. It collects a rich array of information such as individual characteristics, social backgrounds, economic status, religions, personal opinions and attitudes toward some specific topics, etc. The original sample includes 4,528 households and all the full-age members in each household. In later years refreshment samples were added. For example, in June 1990, 2,179 households from former German Democratic Republic (GDR) were included immediately after the reunion. In 1998, a refreshment sample of 1,067 households was added and in year 2000, 6,052 additional households were added. All the information is collected and stored in separate sub-files for each year according to different topics.[26]

[26] Because of this decomposed data structure, both cross sectional and longitudinal analysis requires a matching and merging process.

The longitudinal data constructed out to the SOEP from 1994 to 2004 and is an unbalanced panel. Sample members enter in different years and have different observations from 2 to 11 years.[27]

In the household questionnaire, the question about what kind of savings or investment securities the household owns is asked every year. Life insurance holding is one of the options. I extract this information and use it as the key dependent variable. One drawback is that I can only observe whether or not the sample member has life insurance instead of the detailed contract information such like face and cash value of the policies, premium paid, benefit received, etc. Therefore, I am measuring the continuous change between 0 and 1, or the probability of owning life insurance instead of the quantity purchased. The variable of unemployment is the registered unemployment status at the unemployment office. Compared with the figures from German Federal Statistics Office, it is consistently lower as it omits the unreported unemployment.

Creating the marginal tax rate is a complicated process because of the complex German tax laws. I adopt the same set of simplifying assumptions as did Schwarze (1995): all married persons file jointly; all filing units take the standard deduction; no filing unit itemizes; when no standard deduction exists the allowance is ignored; average national insurance contribution rates for old age pensions, health insurance and unemployment insurance apply to all employees. I incorporate all the changes of tax laws from Year 1994 to Year 2004 into tax functions, from which the key explanatory variable - marginal tax is generated. [28] The relationship between marginal tax rate and the pre-tax income are plotted[29] in Figure 2.1. As it roughly shows, the higher pre-tax income is associated with higher marginal tax rate.

The final longitudinal data consists of 132,144 observations across the period of 1994 to 2004. I restrict the sample to the population between 20 and 65 years old as the incentives to start life insurance contracts after retirement age is

[27] The data structure is as followed: 970 in 1995; 498 in 1996; 443 in 1997; 1800 in 1998; 501 in 1999; 10216 in 2000;674 in 2001; 2774 in 2002; 665 in 2003, 579 in 2004.
[28] Marginal tax rate is defined as the first derivative of the tax function.
[29] I use Year 2004 data only.

low due to the high premium. Table 2.1 summarizes the main variables I use in empirical analysis. The unadjusted mean of life insurance holdings is around 65 percent. The mean age is 44 year old. The average education level is around 12 year. Around 5 percent of the sample report that they are still in education or training. Due to the matching process of household and personal level information, the sample has around 74 percent married individuals.[30] Roughly 9 percent are registered as unemployed at the unemployment office. Seven percent are self-employed and 18 percent are working as civil servants. The average marginal tax rate is about 22 percent. Forty-five percent reported being a home owner and 28 percent reported having mortgage.[31] The average reported satisfaction with self health is about 6.7 out of an 11 point scale. The average risk preference is around 4.5 out of an 11 point scale.

I calculate the transition probability of life insurance holdings. This transition probability is unconditional on any explanatory variables. Across the 11-year period, the probability is around 22 percent for becoming the owner of a life insurance contract and 13 percent for cancelling the contracts. It roughly presents a dynamic picture of life insurance demand in the longitudinal data.

To present a more detailed picture on life insurance holdings, I group the sample members according to the labour force status, family structure and income quartiles. Figure 2.2 provides the comparison between the unemployed and employed. I see a clear gap between these two groups across the time. For the employed, the life insurance holdings stay quite stable and above 55 percent. For the unemployed, it is not only much lower but also decreasing slightly across time, from 52 percent in Year 1994 to 45 percent in Year 2004. Does it mean that the unemployed are worse off under Harz labour market reform[32]? The welfare

[30] The household questionnaire is mainly directed to the head of the household or someone who can stand for the head. Therefore, the variable life insurance is at the household level. I merge the household and personal level information in such a way that only the person reported as the head or the partner are included.

[31] Some banks require that the individual should own life insurance when applying for mortgages due to the liquidity concern.

[32] Harz I and II began to take effect in Year 2003, and Harz III in Year 2004.

change of those unemployed people is naturally an interesting topic worthy of political attention.

Bequest motives could be roughly illustrated by Figure 2.3 and 2.4. Figure 2.3 shows the comparison of life insurance holdings between the married and not married individuals. The married individuals appear to be concerned about their spouse's living standard. The gap between percent of life insurance holdings among the married and not married is around 20 percent. Figure 2.4 shows the difference of life insurance holdings between households with children and those without children. Despite the slightly upward trend for the households with children, the rates are quite stable across the time. The gap between these two groups is about 10 percent. Apparently, the parents are altruistic and concerned about the welfare of their descendents.

Figure 2.5 displays the pattern of life insurance holdings across the 4 income quartiles. Clearly life insurance is a normal good. Higher income is related to higher life insurance holdings across years. The largest gap exists between those in the lowest quartile and those in the second quartile. The percent of having life insurance is nearly doubled in the later group. Both of these two group show slightly decreasing holding rates across time. The increasing number of life insurance contracts signed mainly comes from the groups that have higher than average income. Moreover, the story of tax advantages might help explain the higher demand among the high-income earners. As their marginal tax rates are higher, the motives to buy life insurance to utilize the tax deductibles are higher.

IV. Reduced-Form Estimations and Empirical Findings

As the theoretical model implies, demand for life insurance is a function of the implicit price of the insurance, household's risk aversion, the accumulated wealth, labour force participation, the household's weights on bequests, and tax parameters. In the empirical analysis I use reduced form estimation.

I start with a baseline equation:

(1) $y_{it} = a + x_{it}\beta + \upsilon_{it}, i = 1,....,N; t = 1,....,T$

where y_{it} and x_{it} are observations for the ith individual at time t and β is a vector of coefficients. Specifically, in my empirical analysis, y_{it} is the dummy of life insurance holding where 1 indicates having life insurance and 0 having none). x_{it} is a set of explanatory variables including sex, age and age squared, education level, education or training status, marital status, number of children, housewife status, house owner, mortgage holder, whether unemployed or self-employed, whether working as civil servant, net labour income and labour income squared, and marginal tax rate. I also add the subjective measure of satisfaction of self-health to test the existence of adverse selection into life insurance holdings. If adverse selection is present, then I would expect that the person who is less satisfactory with one's own health would be more likely to buy life insurance. In reality, the insurance company usually checks the applicant's health status carefully. Therefore adverse selection may not be serious. I control the cohorts, nationalities, and each of the 16 states as well. In the empirical estimate for Year 2004 cross sectional data, the self reported risk attitude is added to control for individual heterogeneity.

The bequest motives are modelled in marital status, number of children and the interaction term between the dummy of married and the dummy of having children at home. Unemployment and housewife status capture the effect of exit from the labour force. The self-employed and civil servants are subjected to different treatment from German tax and the public pension system: the self-employed are not obliged to contribute to the public pension system but must provide for their own retirement income and survivor's benefits; civil servants with tenure do not need to contribute to the pensions but receive quite generous survivor benefits. Tax preference is modelled by the marginal tax rate variable.

In the baseline equation, υ_{it} is the residual with mean zero and variance σ_{ν}. It is assumed to be uncorrelated with x, and there is no autocorrelation. This is the easiest and most convenient assumption about the error structure. Pooled OLS is used to estimate the simplified model. In doing so, I pool cross-sectional time series, which incorporates both the cross-sectional effect of x on life insurance demand as well as the time series effect. All the individuals in the data are

characterised by the same regression equation (1) at all points in time. However, it turns out the pooled cross sectional time series is quite difficult to estimate. Pooled OLS in most cases fails to perform this task. As Hicks (1994) points out, the error terms from pooled OLS regression tend to be temporally autoregressive, cross-sectional heteroscedastic and correlated and may conceal unit and time effects. Table 2.2 displays the Wooldridge test for serial correlation from the regression of the first differences and yields an F statistics of 374.59. Therefore the OLS regression leads to inefficient estimates.

There are three ways the current econometric literature dealing with autocorrelation. The first one is a static model, in which autocorrelation is regarded as a nuisance in the residuals and has to be corrected. The second one is a dynamic model, in which autocorrelation is treated as persistency in the dependent variable and is captured by modelling an autoregressive process including a lagged dependent variable. In the third approach, autocorrelation in pooled time series is regarded as resulting from unit roots in the single series and is can be corrected by differencing the series[33].

I use the static approach where the error structure follows a first-order autoregressive or AR(1) process:

(2) $y_{it} = a + x_{it}\beta + \mu_i + e_{it}, i = 1,...., N; t = 1,...., T$

where

$$e_{it} = \rho e_{i,t-1} + z_{it}$$

and where the absolute value of the autocorrelation parameter, ρ, is less than 1 and z_{it} is independent and identically distributed with zero mean and variance σ_z^2. The disturbances are modelled as a first-order autoregressive process.[34] In this random effects model, μ_i are independent of x_{it} and assumed to be realizations of an i.i.d process with mean zero and variance σ_z^2 [35]. The Breusch-Pagan lagrange multiplier test (see Table 2.2) on the absence of random

[33] See Wooldridge (2003) Ch.13.
[34] This can be implemented by Stata program with a random effects GLS regression.
[35] I assume there is no between group autocorrelation.

effect yields a test statistic of 1883.49, which exceeds by far the 95 percent critical value for a chi-squared statistic with one degree of freedom, 3.84. Without question the random effect model fits better than the pooled OLS with the data. Moreover, I perform the Hausman specification test on the model fit of the fixed effect model versus the random effect model. The null hypothesis that the fixed effect model fits better is rejected with the highly significant test statistic of 727.87.

As the dependent variable is a dummy which takes value of 0 and 1, the linear probability model is not the most satisfactory one as it can generate a probability that is either greater than one or less than zero. It also implies a constant marginal effect of each explanatory variable that appears in its original form. To avoid the linear probability model's problem, I proceed to explore the panel probit model that also takes into account the autoregressive error structure. The model can be written as follows:

$$(3) \quad y_{it}^{*} = a + x_{it}\beta + v_{it}, i = 1,...., N; t = 1,...., T$$

$$y_{it} = \begin{cases} 0, y_{it}^{*} \leq 0 \\ 1, y_{it}^{*} > 0 \end{cases}$$

The disturbances v_{it} are normally distributed with a T×T positive definite covariance matrix Σ. The typical element of Σ is denoted σ_{ts}. In the empirical estimation, however, I do not restrict the error structure to be AR(1) process but instead as an unstructured one. x_{it} are assumed throughout to be strictly exogenous, which implies that Cov[x_{it},v_{js}] = 0 across all individuals i and j and all periods t and s. This rule out state persistence or the presence of lagged dependent variables on the right hand side.

The results of the random effect model with AR (1) errors are displayed in Table 2.2 and the results of the panel probit model with marginal effects are displayed in Table 2.3. Both models yield consistent results except that most of the estimates from the panel probit model are slightly larger in magnitude.

The probability of life insurance holdings depends nonlinearly on age and income as captured by the negative coefficients on their squared terms.

Graphically, it exhibits a hump shape in age and income. This is consistent with the previous research results on the relationship between life insurance holdings and age and income. I control the cohort effects in both models as well. The results indicate that the younger cohorts are less likely to own life insurance, whereas the cohorts born between 1940s' and 1960s' are more likely to purchase life insurance. In other words, individuals age 30 to 60 during the study period have a greater tendency to purchase life insurance after controlling for other factors.

The bequest motives are clear in life insurance holdings. Married couples are more likely to own life insurance than single people. The probability increases when married couple have children. It is about 14 percent more likely that a married couple with a child has life insurance than its single counterpart. The number of children appears to have a slightly negative impact on life insurance holdings. The presence of a dependent child at home positively affects the bequest motive. More children means the parents face tighter financial constraints but also unofficially guarantees care when parents get older.

Life insurance holdings are significantly lower among the unemployed: approximately 2.3 percent and 2.4 percent in the random effects model and panel probit model, respectively. Housewives are significantly less likely insured. Those who are still in studying or training and are not yet in the labour force are less likely to have life insurance, probably due to the decreased financial ability to buy life insurance. The self-employed are about 5 percent more likely to have life insurance. As the self-employed must provide their own pensions and survivor benefits, they are more prone to purchasing life insurance. I did not find any significant effect in either model on owning life insurance among civil servants, as they usually enjoy generous survivor benefits provided by the government.

Tax incentives are highlighted by the significant coefficient of the marginal tax rate. Marginal effects are 23.9 and 22.5 percent in RE and panel probit models, respectively. In other words, if a person is subject to an average marginal tax rate of 22 percent (each extra dollar earned is taxed at a rate of 22 percent), then the probability is 5 percent higher than if he were subject to zero tax.

As the life insurance payments were enjoying tax-free treatment, a higher marginal tax rate provides a stronger incentive to take advantage of tax preferences towards life insurance. Anticipation of the tax benefit abolishment on contracts signed after 2004 speed up people's decision to own life insurance. In later analysis, I investigate this program anticipation effect two years prior to the implementation of new rules.

I find that home owners, especially those with mortgages from banks are 5 to 6 percent more likely to buy life insurance. As mentioned before, some banks require that applicants have life insurance contracts before approval of loans in case of default.

Another interesting variable is the subjective measure of satisfaction with self health. It has small but significant positive effects in both models. People who are more satisfied with their own health are more likely to have life insurance. This is closely related to the operating mechanism of insurance companies so that they check the health status of applicants very carefully upon signing contracts. People who are in bad health situation are less likely to get approved. There is no adverse selection into life insurance, as previous literature indicates.

I control for cohorts, nationality and each of the 16 states in both models. Individuals with German nationality have a stronger tendency to buy life insurance than foreigners living in Germany, such as those from Turkey, Italy, Greek and other countries. Controlling for each of the 16 states shows that, in general, West Germans are less likely to purchase life insurance than their East German counterparts.

Table 2.4 reports results of a probit model specifically run for the 2004 sample in order to examine the relevance of the subjective measure of risk preferences in life insurance demand. This self-reported risk variable appears to be statistically insignificant in explaining life insurance holdings. However, using only the 2004 sample has a disadvantage in that it is cross sectional data and therefore lacks personal variation across the time used for properly identifying the effect of risk attitudes. I construct another dataset that contains an observation period from 2002 to 2004 and simply traces back the risk attitude, as it is unusual

that the individual changes risk attitudes frequently during this three-year period. I then find a small but significant positive effect (see Table 2.4). This self-reported risk attitude captures the individual heterogeneity across time. Some previous studies find that individuals reporting higher risk preferences are more likely to take risky actions such as portfolio choice, self-employment, mobility, etc. Therefore, it could be positively related to behaviour in holding life insurance.

Life insurance has certain tax advantages as shown by the effect of the marginal tax rate on the demand for life insurance. However, tax benefits are being abolished on contracts signed after Dec.31, 2004. The German government started to discuss this issue early in 2003. How would the public respond to the potential change in tax policies to life insurance? Would those without life insurance start shopping around to catch the bus? I run a separate panel probit model of life insurance demand based on data from 2002 to 2004 to approximate the program anticipation effect. Like before, I control for self-reported risk attitudes. I find that purchasing life insurance increases in 2003 and 2004 compared with 2002, when policy discussions had not yet started. It confirms the anticipation effect of the policy change on people's behaviour. In 2003 the increase was about 17 percent, and in 2004 18 percent compared with 2002[36]. Both appear to be statistically significant at the 1 percent level. To evaluate the program more precisely, it is important to obtain the data after 2004. This outside of the data range and would be an interesting research topic once the data of a longer post-ante period is available.

V. Conclusion

As life insurance is an important component of a household's private savings in Germany, understanding the decision to hold life insurance sheds some light on life-cycle savings behaviour. This chapter studies the demand for life insurance in Germany by examining the driving forces behind the demand behaviour.

[36] The regression results are upon the request.

The theoretic model is a life cycle model that incorporates the specific features of the German tax and pension system. I test theoretical predictions of the impacts of bequest motives, labour market participation, and tax advantages on the demand for life insurance using German SOEP data. In addition, I examine the program anticipation effect of abolishment of the tax benefit. And I utilize self-reported risk attitudes to capture individual heterogeneity across time as well.

The panel probit model, which takes care of the nature of dummy dependent variables, produces slightly larger marginal effects than the random effects model. In general, I find that bequest motives are quite strong among married families with children. It is around 14 percent more than their single counterparts without any children. The unemployed are consistently less likely to be insured in the samples. The same is true for housewives, people in still in school or training and low-income families. Higher marginal tax rates provide individual incentives to buy life insurance. The anticipation of the abolishment of tax benefits on life insurance increases the demand prior to the real change. Moreover, I find self-reported risk attitudes turn out to be positively related with life insurance holdings.

Although the data allowed us only to examine the possibility of holding life insurance, it could be interesting to obtain a dataset which would enables us to examine the quantity demanded. This would require a different dataset that records more detailed life insurance holdings such as the type of the contracts, the monthly or yearly premium, amount insured for, etc. Further research could also utilize the GSOEP data released after 2005 in order to perform program evaluations on the policy change of tax treatments on life insurance.

Figure 2.1: Relationship between Marginal Tax Rate and Income
Year 2004 Sample

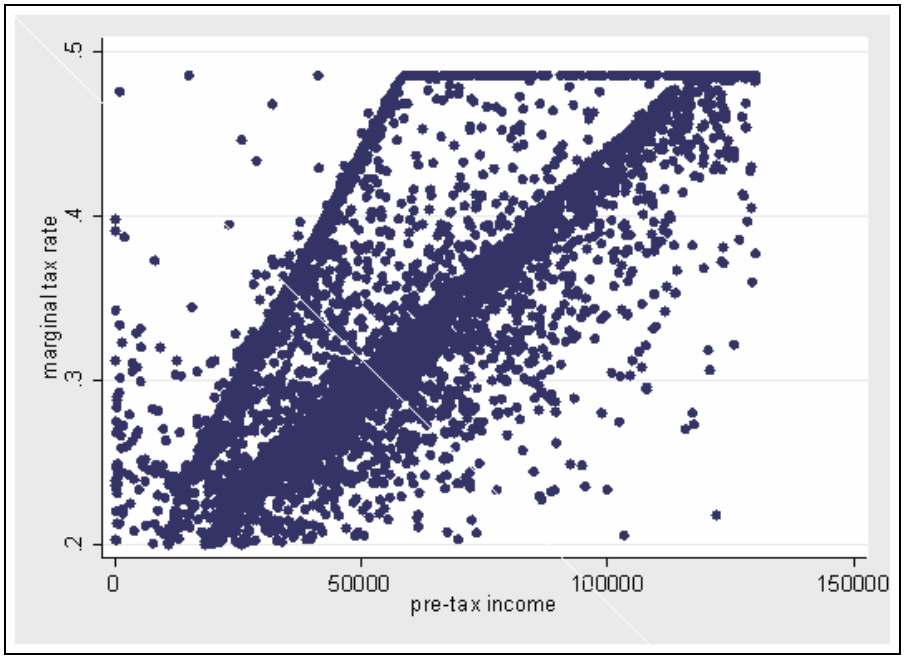


Figure 2.2: Comparison of LI Holdings between the Employed and Unemployed
1994-2004

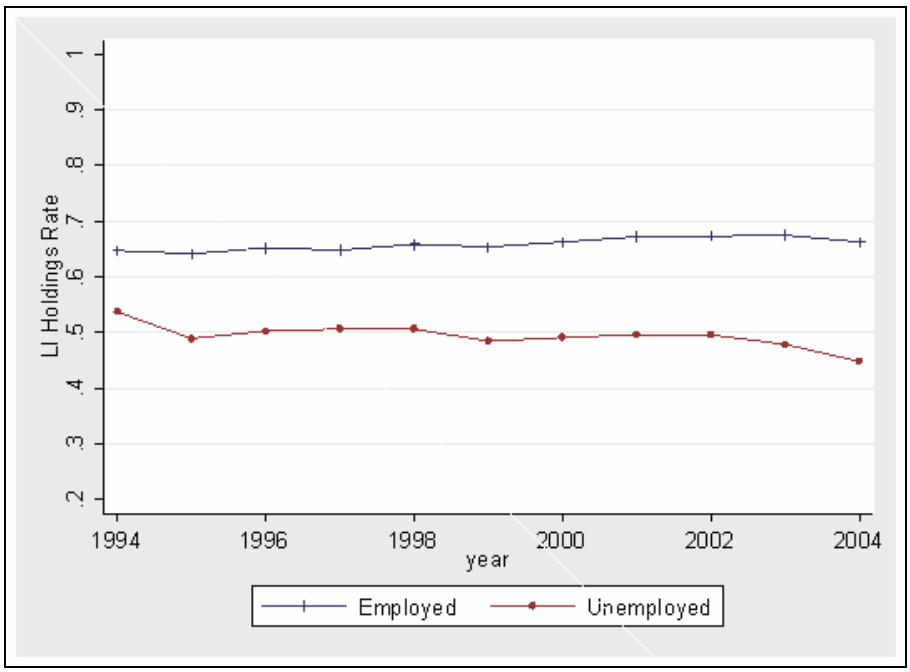


Figure 2.3: Comparison of LI Holdings between the Married and not Married

1994-2004

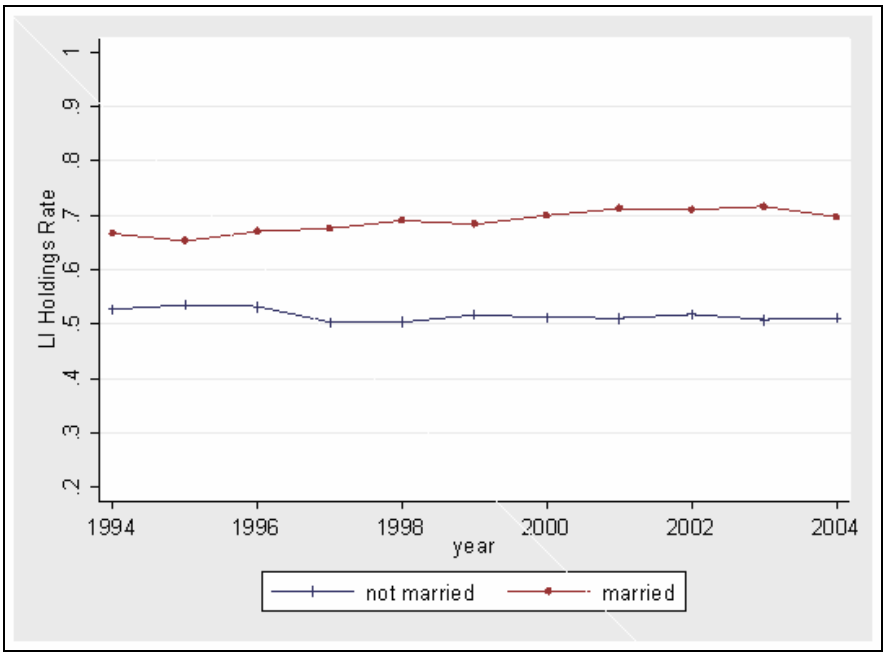


Figure 2.4: Comparison of LI Holdings between Households with & without Children

1994-2004

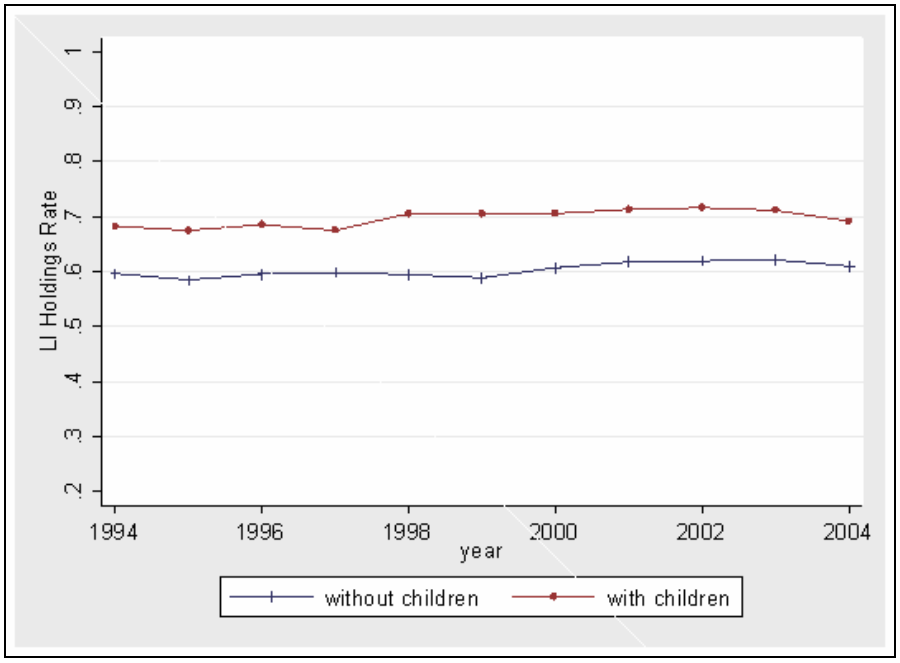


Figure 2.5: Comparison of LI Holdings across 4 Income Quartiles

1994-2004

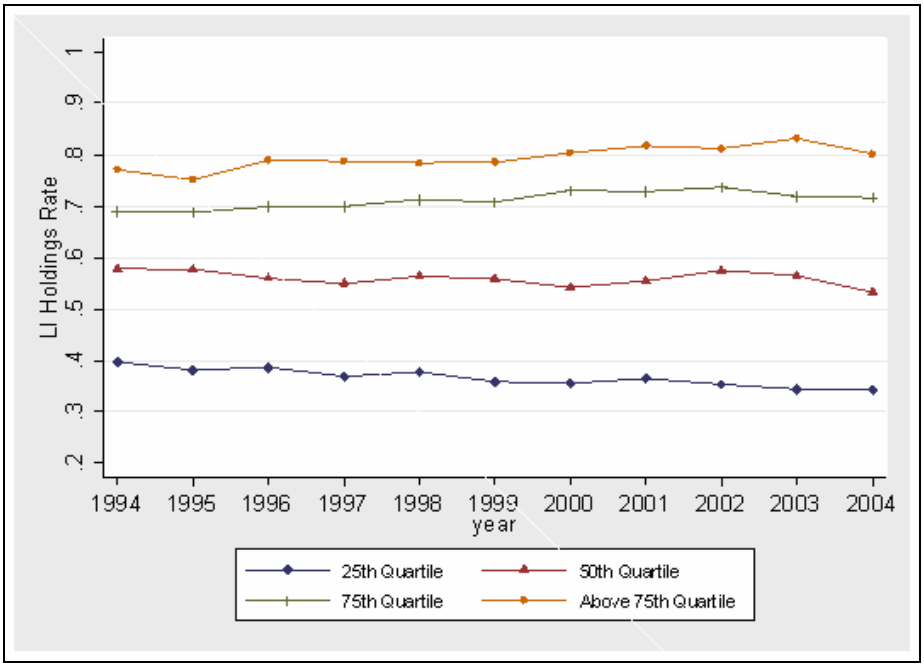


Table 2.1: Summary Statistics of Main Variables

Variable	Mean	Std. Dev.	Min	Max
Life Insurance Holding	0,646	0,478	0	1
Age	43,978	11,750	20	65
Male	0,476	0,499	0	1
Education level	11,913	2,617	7	18
Still in education or training	0,049	0,216	0	1
Married	0,741	0,438	0	1
Number of children	0,735	1,000	0	10
Housewife	0,365	0,482	0	1
Unemployed	0,087	0,282	0	1
Self-employed	0,068	0,251	0	1
Civil Servant	0,179	0,384	0	1
House owner	0,450	0,497	0	1
Mortgage	0,279	0,448	0	1
Net labour income	19731,69	19687,15	0	146350
Marginal tax rate	0,222	0,143	0	0,53
Satisfaction with own health	6,713	2,147	0	10
Risk Attitude from 0 to 10	4,528	2,278	0	10
Nationality German	0,883	0,322	0	1
Nationality Turkey	0,040	0,197	0	1
Nationality Italy	0,019	0,137	0	1
Nationality Greek	0,012	0,111	0	1
Nationality Ex-Yugoslavia	0,012	0,109	0	1
Cohort born in 1930s'	0,107	0,309	0	1
Cohort born in 1940s'	0,189	0,392	0	1
Cohort born in 1950s'	0,232	0,422	0	1
Cohort born in 1960s'	0,256	0,437	0	1
Cohort born in 1970s'	0,099	0,299	0	1
Cohort born in 1980s'	0,008	0,090	0	1

Note: Sample period is from 1994 to 2004.
The variable 'Net labour income' refers to the individual labour income. I exclude the highest one income percentile.

Table 2.2: Random Effect model of Life Insurance Demand with AR(1) Disturbances

Explanatory Variables	Coefficient
Age	0,017***
Age squared	-0,000***
Male	-0,011**
Education level	0,005***
Still in education or training	-0,023***
Married	0,121***
Number of children	-0,007***
Married*children	0,021***
Housewife	-0,010**
Unemployed	-0,023***
Self-employed	0,047***
Civil Servant	0,005
House owner	0,016***
Having mortgage	0,036***
Net labour income	1.81e-06***
Net labour income squared	-6.48e-12***
Marginal tax rate	0,239***
Satisfaction with own health	0,002***
Constant	-0,078**
5 biggest nationalities	Controlled
Cohorts	Controlled
16 states	controlled
# of observations	132144
Wald Chi2	5894,37
Breusch Pagan	1883,49
Baltagi-Wu LBI	1,97
Hausman test	727,87
Woodridge test	374,586

Source: German Socio-economic Panel (1994-2004)

Note: Breusch Pagan lagrange multiplier test on absence of random effects: $H0 : var(\mu_i) = 0$. Baltagi-Wu LBI is the Baltagi-Wu (1999) locally best invariant test statistic from the AR(1) model, $H0 : \rho_r = 0$; ρ_r is estimated error autocorrelation coefficient. Hausman test is Hausman's (1978) specification test of the appropriateness of fixed effect model. Wooldridge test for serial correlation from the regression of the first-differences variables, H0 : no first-order autocorrelation.
The other four biggest nationalities in the data are Turkey, Greece, Italy, Ex-Yugoslavia. *** means significant at the level of 1% .** means significant at the level of 5%. * means significant at the level of 10%

Table 2.3: Panel Probit model of Life Insurance Demand with Correlated Errors

Explanatory Variables	Coefficients	Marg. effect
Age	0,053***	0,020***
Age squared	-0,001***	-0,000***
Male	-0,033**	-0,012**
Education level	0,016***	0,006***
Still in education or training	-0,059***	-0,022***
Married	0,327***	0,124***
Number of children	-0,021**	-0,008**
Married*children	0,077***	0,028***
Housewife	-0,029**	-0,011**
Unemployed	-0,064***	-0,024***
Self-employed	0,141***	0,051***
Civil Servant	0,018	0,007
House owner	0,048***	0,018***
Having mortgage	0,116***	0,043***
Net labour income	4,92E-06***	1.83e-06***
Net labour income ^{2}	-1,09E-11	-4.05e-12
Marginal tax rate	0,606***	0,225***
Satisfaction with own health	0,006***	0,002***
Constant	-1,728***	---
5 biggest nationalities	controlled	
Cohorts	controlled	
16 states	controlled	
# of observations	132144	
Wald chi2	4406.04	

Source: German Socio-economic Panel (1994-2004).

Note: Marginal effects for dummy variables are calculated for discrete change from 0 to 1.
*** means significant at the level of 1%
** means significant at the level of 5%
* means significant at the level of 10%

Table 2.4: Probit Model of Life Insurance Demand with Risk Measure

Explanatory Variables	2004 Sample (Probit)		2002-2004 Sample (Panel Probit)	
	Coefficients	**Marg. effect**	**Coefficients**	**Marg. effect**
Age	0,052***	0,019***	0,061***	0,022***
Age squared	-0,001***	-0,000***	-0,001***	-0,000***
Male	-0,047*	-0,017*	-0,021	-0,007
Education level	0,011**	0,004**	0,013***	0,005***
Still in education or training	-0,071	-0,026	-0,083***	-0,030***
Married	0,393***	0,148***	0,411***	0,153***
Number of children	-0,047**	-0,017**	-0,065***	-0,023***
Married*children	0,042	0,015	0,114***	0,041***
Housewife	-0,077**	-0,028**	0,007	0,003
Unemployed	-0,178***	-0,067***	-0,063***	-0,023***
Self-employed	0,180***	0,064***	0,101***	0,036***
Civil Servant	0,007	0,003	-0,013	-0,005
House owner	0,219***	0,080***	0,241***	0,087***
Having mortgage	0,085**	0,031**	0,051**	0,018**
Net labour income	1,52E-06	5.59e-07	5,89E-06***	2,12E-06***
Net labour income ^{2}	7,66E-12	2.81e-12	-3,38E-11***	-1,22E-11***
Marginal tax rate	1,436***	0,527***	0,703***	0,253***
Satisfaction with own health	0,010**	0,004**	0,003	0,001
Risk attitude in 2004 sample	*0,008*	*0,003*	---	---
Risk attitude in 2002-2004 sample	---	---	*0,018***	*0,007***
Year 2003	---	---	0,49281***	0,17381***
Year2004	---	---	0,52081***	0,181***
Constant	-1,877***	---	-1,857***	---
5 biggest nationalities	controlled			
Cohorts	--			
16 states	controlled			
# of observations	14205		45085	
Test Statistics	Log likelihood: -8197.7024		Wald chi2: 2074.35	

Source: German Socio-economic Panel (2002-2004).

Note: Marginal effects for dummy variables are calculated for discrete change from 0 to 1.
*** means significant at the level of 1%. ** means significant at the level of 5%. *means significant at the level of 10%

Reference

Alan J. Auerbach, Laurence J. Kotlikoff, (1991) *"The Adequacy of Life Insurance Purchases."* Journal of Financial Intermediation, Vol. 1, No. 3, pp. 215-241,.

Baltagi, Badi H. and Wu, Ping X (1999), *"Unequally Spaced Panel Data Regressions. with AR (1) Disturbances."* Econometric Theory, Vol. 15, pp. 814-23.

Bernheim, B. Douglas, Lorenzo Forni, Jagadeesh Gokhale and Lawrence J. Kotlikoff. (2003) *"The Mismatch between Life Insurance Holdings and Financial Vulnerabilities: Evidence from the Health and Retirement Study,"* American Economic Review, v93 (1, Mar), 354-365.

Bernheim, D. (1991) *"How Strong Are Bequest Motives? Evidence Based on Estimates of The Demand for Life Insurance and Annuities",* Journal of Political Economy, pp.899–927.

Bonin et al. (2006) *"Cross-Sectional Earnings Risk and Occupational Sorting: The Role of Risk Attitudes",* IZA working paper, Jan.2006 (forthcoming in: Labour Economics, 2007)

Dohmen et al. (2005) *"Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey",* IZA discussion paper, 1730.

Browne, Mark J.; Kim, Kihong, (1993) *"An International Analysis of Life Insurance Demand".* Journal of Risk and Insurance: Dec, issue

Campbell, R.A. (1980) *"The Demand for Life Insurance: An Application of the Economics of Uncertainty",* The Journal of Finance, Vol. 35, No. 5, pp. 1155-1172.

Chang, Fwu-Ranq (2004) *"Life Insurance, Precautionary Saving and Contingent Bequest"* Mathematical Social Sciences, Vol. 48, pp. 55-67

Drukker, D. M. (2003) *"Testing for Serial Correlation in Linear Panel-data Models".* Stata Journal 3: 168–177.

Fischer, S. *A Life Cycle Model of Life Insurance Purchases.* International Economic Review *14*: 1973, pp.132–152.

Gesamtverband der Deutschen Versicherungswirschaft, *Statistical Yearbook of German Insurance 2003*

Gesamtverband der Deutschen Versicherungswirschaft, *Statistical Yearbook of German Insurance 2004*

Gokhale, Jagadeesh and L. J. Kotlikoff (2002), *"The Adequacy of Life Insurance", Research Dialogue,* TIAA-CREF INSTITUTE, tiaa-crefinstitute.org, issue no. 72 July 2002.

Green, William, (2002), *"Convenient Estimators for the Panel Probit Model: Further Results "*, Empirical Economics, Volume 29, No.1, January 2004.

Hong, J. H. & R´ıos-Rull, J.-V. (2004). *"Life Insurance and Household Consumption"*. FRBP Working Paper 04–10..

Jappelli, Tullio & Pistaferri, Luigi, (2003). *"Tax Incentives and the Demand for Life Insurance: Evidence from Italy,"* Journal of Public Economics, vol. 87(7-8), pages 1779-1799, August.

Ostaszewski, K. (2003*), "Is Life Insurance a Human Capital Derivatives Business?"* Journal of Insurance Issues, 2003, 26, 1, pp. 1-14.

Pauly, Mark V. , Withers , Kate H.; Subramanian-Viswana, Krupa; Lemaire, Jean; Hershey, John C. (2003**)** *"Price Elasticity of Demand for Term Life Insurance and Adverse Selection"*. NBER Working Paper No. 9925.

Richter, Andreas and Ruß, Jochen (2001), *"Tax Arbitrage in the German Insurance Market"*. Blätter der Deutschen Gesellschaft für Versicherungsmathematik 25 (2003): 659-672

Showers, Vince E. and Shotick, Joyce A. (1994) *"The Effects of Household Characteristics on Demand for Insurance: A Tobit Analysis"*. The Journal of Risk and Insurance, Vol. 61, No. 3, pp. 492-502

Smith, Michael L. and Stephen A. Buser (1983), *"Life Insurance in a Portfolio Context"*, Mathematics and Economics, 2, 1983; 147-57

Truett, Dale and Truett, Lila, (1990) *"The Demand for Life Insurance in Mexico and the United States: A Comparative Study"* The Journal of Risk and Insurance, Vol. 57, No. 2, pp. 321-328

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press

Winter, Joachim and Walliser, Jan. (1998) *"Tax Incentives, Bequest Motives and the Demand for Life Insurance: Evidence from Germany"*. Discussion Paper 99-28, Sonderforschungsbereich 504, Universität Mannheim.

Yaari, M.E.: *"Uncertain Lifetime, Life Insurance and the Theory of the Consumer"*. Review of Economic Studies, 32, 1965, pp.137-150.

Chapter Three

Risk Selection and Inefficient Provision of Supplementary Health Insurance in Germany

The rising health expenditure and increasingly constrained public resources in Germany has led to several reform packages in recent decades. Due to the cuts in benefits and coverage in the statutory health insurance market, more and more publicly-insured members purchase supplementary health insurance through the private market in order to cover the unforeseen risks. However, using the data from German Socio-Economic Panel Studies, I find clear evidence of risk selection and inefficient provision of supplementary health insurance in the private health insurance market. Controlling for the measurement problem of self-reported health status in my empirical models, I find that poorer health status is negatively related with the purchase of supplementary health insurance, especially those plans covering hospital stays. The findings in this chapter have important implications for the current debate on patient's data protection if the government aims to improve the efficiency of the private health insurance market in providing supplementary coverage.

I. Introduction

Health care expenditures have been increasing and slowly draining public resources in many countries. In the case of Germany, about 158 billion Euro were spent on health in 1992. The figure rose to 201 billion in 1998 and 234 billion in 2004. Health expenditures per inhabitant in Germany rose from 1960 Euro in 1992 to 2840 Euro in 2004[37]. Policy makers and health economist have been searching for suitable alternatives to finance health care.

An observed trend of health care reforms in industrialised countries is introducing a mixed health insurance system, where public health care requires patient co-payments for services. This is well illustrated by Zweifel and Breyer (1997). A mixed public/private health insurance system is also modelled in some recent research.[38] Several OECD countries' governments consider supplementary health insurance as one of the primary instruments for limiting statutory financing of health care.[39]

A very early study of the demand for supplementary health insurance can be found in Newhouse et al. (1977). Their paper mainly focused on the model incorporating tax subsidies.[40] Seidman (1978) developed a criterion for the optimal treatment of supplementary health insurance under income-related major-risk (catastrophic) national health insurance (MR-NHI) and used it to derive the optimal treatment under an MR-NHI proposal that may be enacted in U.S.A. Petretto (1999) has developed a theoretical framework of National Health Service where compulsory social insurance covers a package of essentials and supplementary private policy tops up the remaining services. In his paper the model was solved by backward induction with a three stage maximization process, and the conditions for optimal rates for social insurance coverage and private coinsurance are analysed. In equilibrium, the private insurance contract signed by each individual includes a coinsurance rate which is a function of social payroll

[37] Source: the Federal Health Monitoring system, http://www.gbe-bund.de
[38] For example see Blomqvist and Johansson (1997).
[39] Organisation for Economic Co-operation and Development. *Private health insurance in OECD countries.* The OECD Health Project. OECD, 2004.
[40] They show that the demand for supplementing outpatient services will be small unless the tax subsidy of insurance is continued.

tax and social insurance rate. The optimal social insurance coverage positively depends on the distributional characteristics of health services and also the individual's gains from risk-sharing, that is, out-of-pocket health expenditures.

In many industrialised countries with a National Health Service System such as Germany, the principle of funding statutory health care is based the concept of solidarity - the belief that society is responsible for the well-being of its members. This means the contributions are made according to ability to pay and people receive benefits according their health care needs. However, as pointed out by van de Ven et. al (2006), in most voluntary supplementary health insurance markets these institutional and regulatory arrangements, adopted by governments with the aim of guaranteeing solidarity, are formally absent. In the long run, the absence of these legal constraints may induce insurers to risk-rate premiums in order to attract the better risks and thereby increase their competitiveness and profits. In Germany, health insurance companies can quite easily identify different risks through full access to patients' data. Meanwhile, the insurers have no obligation to enrol any applicant who needs some sort of supplementary health insurance. The optimal strategy for profit-maximizing insurers is to design selection techniques to insure good risks and exclude bad risks. Risk selection techniques include selective underwriting, benefits package design, selective advertising, denial of coverage, exclusion of pre-existing medical conditions, differential waiting periods, termination of contract, etc.

The purpose of this chapter is two-fold. First of all, I attempt to confirm the existence of risk selection in the supplementary health insurance market with the empirical data. Second, econometrically I aim to control the endogeneity problem of the self-reported health status to obtain more accurate estimates of its effect on supplementary health insurance holding.

There are few empirical studies concerning the demand for supplementary health insurance in Germany. Christoph (2002) used German Social-Economic Panel (GSOEP) data to study the demand for supplementary private health insurance among compulsory members. They find that better health status is related to a higher probability of having supplementary health insurance. Their

argument is that people with bad health may face more restrictions when applying for supplementary health insurance. Their bad health results in either higher premiums or a rejection of their application. However, their model doesn't take into account measurement error in health status and therefore the estimated effect of health is biased.

Health status is an important variable in studying the demand for health insurance. As indicated in the health economics literature, adverse selection is an important phenomenon of the health insurance market. Adverse selection occurs since high risks tend to consume more insurance than low risks in a market with asymmetric information and favours the buyer.[41] Health economists quite often use the individual health status in empirical models to test for the existence of adverse selection in the insurance market.[42] If good health is related to the purchase of less health insurance while bad health is to the higher amounts of insurance, all else being equal, adverse selection is present.

In survey questions, the respondents are often asked to rate their own health. However, the direct application of self-assessed health status is usually questionable due to measurement error. An early analysis by Angel and Gronfein (1988) demonstrated that the social construction of such subjective information makes its use in comparative analysis problematic since both its accuracy and the outcome for which it is employed as a predictor are influenced by the respondent's culture and social location. A recent study by Crossley and Kennedy (2002) on the reliability of self-assessed health shows that there is substantial error in rating and this error is highly correlated with observable variable such as age, gender and income. Therefore, simply employing the self-reported health status from German Social-Economic Panel would result in an incorrect estimate of the effect of health status on social outcomes of interest.

From German Social-Economic Panel data I select wave 2002 and wave 2004 and construct a balanced panel. Using the predicted health status instead of the self-reported health, I found that worse health status is negatively related to

[41] See for example Pauley (1974), Brown (1992).
[42] See Brown (1992). He used self-reported health status in National Medical Care Expenditures Survey to test for adverse selection in the individual medical expense insurance market.

the holdings of supplementary health insurance and the amount of purchase. While this study did not find any evidence of adverse selection in the private health insurance market in terms of providing supplementary coverage, it does confirm the existence of risk selection in this market. Furthermore I examine the holdings of supplementary coverage for hospital stays as incurring hospital care can greatly impact household financial status. I found clear evidence of risk selection as well.

The rest of this chapter is organized as follows. Part two provides a brief on the institutional background in Germany. Part three presents the empirical models I use in the estimation. Part four describes the dataset I construct out of GSOEP. Part five discusses the empirical results. Conclusion can be found in Part six.

II. Institution Background of German Health Insurance System

In order to understand the driving forces behind the purchase of supplementary health insurance, some background knowledge about the institutional setup in the German health insurance system is important. Statutory health insurance (referred as public health insurance) is a centrepiece of the German welfare state. Approximately 70 million out of a total population of 82 million people are covered by the statutory health insurance. Only a relatively small number of persons- about 300,000 are uninsured. The rest of the population are covered by the private health insurance.[43]

Solidarity is one important principle of the statutory health insurance, and it means the funding of health care and access to it are based on the ability to pay and the need of health care. The contribution to the statutory health care system is proportional to individual's gross income which is under an income ceiling. For example, in 2006 the average contribution rate of statutory health insurance was about 13.4 percent, which is shared equally by the employer and the employee. The upper income ceiling for the contribution to the statutory health care system

[43] Federal Health Monitoring system, http://www.gbe-bund.de

is 42,750 Euro.[44] Private health insurance, on the other hand, charges a premium according to the individual risk and the actual coverage purchased.

Statutory health insurance is characterized by the large volume of interpersonal redistribution, as the contribution and benefits are actually de-linked. It is also a family insurance, which automatically covers the children and a non-employed spouse without any additional charge. They enjoy the same benefits as the paying members. Subject to certain conditions, the statutory system also covers pensioners, the unemployed, trainees and students. The publicly insured members have the opportunity to obtain the additional coverage in the private insurance market.

Since 1980s', many health care reforms have been seen in Germany in an attempt to control soaring costs. In particular, the reform packages in 1997 contained several provisions that intended to control public health expenditures and increase revenues. Cash benefits to sick employees were reduced and the access to treatments at spas was restricted. Higher co-payments were demanded from patients. For example, the co-payment per day of in-patient care increased from 5 DM to 17 DM per day and additionally for a maximum of 14 days per year. And some recent reforms have increased the co-payment per day to 10 Euro (roughly equal to 20 DM) and for a maximum of 28 days. Since 2006, public health insurance no longer covers glasses either. Private health insurance companies provide a variety of supplementary contracts which cover hospital stays that are longer than the maximum set by public insurance funds, dental services and corrective devices that are outside the coverage provided by public insurance funds, as well as a period of travelling abroad, and other kinds of service such as visiting medical practitioners which do not belong to the traditional schools of medicine that is contracted by public insurance funds.

German laws and regulations enable a high degree of transparency in market information. In general the applicant must fill out their health declaration form when filling out the application form for supplementary health insurance in the private health insurance market. They must describe their pre-existing health

[44] http://www.versicherungsnetz.de/

conditions, such as cancer, and these pre-existing conditions are of great concern for the insurer. They are asked how their general health is in the last 5 years, and specifically in last 10 years whether there were some operation or surgery in the hospital and how long the hospital stay was. This is the so-called look-back period. If the applicant received medical advice, recommendations, prescription drugs, diagnosis, or treatment for a health problem during the look-back period, he or she is considered to have a pre-existing condition. The applicant cannot lie about his health history as the private health insurance companies can always access his information through the family doctor or other ways. Then the private health insurance companies will classify the applicant into high or low risk and decide whether to reject the application or adjust the premium according to the individual's risk.

Starting the end of 2006, the political discussion began about whether the patient's data should be protected from health insurance companies and the applicant should have the right to keep his information. If the new policy passed to protect the patients' records, private health insurance cannot detect the type of the applicant as easily. Risk selection will be restricted. It is possible that the companies will have some mechanism to design the contract in such way to attain a separating equilibrium or pooling equilibrium[45] under the new circumstances. However, in this chapter I mainly focus on the risk selection of current market. Further potential problems of adverse selection and moral hazard arising under the new circumstances are outside of the scope of this study.

III. Empirical Model

In this study I am modelling the holding, not the demand, of supplementary health insurance. As mentioned in Part II, private health insurance companies are not obliged to provide supplementary coverage to all applicants. Due to this risk selection, we can only observe those contracts closed by the insurer and the insured in our GSOEP data and we do not have any information

[45] A separating equilibrium is characterized by the low risks and high risks purchasing different contract while a pooling equilibrium is characterized by a subsidy of high risks insurance coverage by low risks.

about whether the individual has applied for the supplementary coverage. In this sense, we actually do not observe the demand but instead just the actual holding status. As the rejection rate shall be greater than zero, there must be excess demand for supplementary health insurance in the private health insurance market.

The following equation models the holdings of supplementary health insurance:

(1a) $y_{it}^{*} = \beta_1 x_{it} + \alpha * h_{it} + \varepsilon_{it}$ $t=1, 2$ *and* $i=1....N$

where y_{it}^{*} is an unobserved latent variable for supplementary health insurance. In our data we observe y_{it} which equals to 1 (having supplementary health insurance) or 0 (having no supplementary health insurance).

$$y_{it} = \begin{cases} 0, y_{it}^{*} \leq 0 \\ 1, y_{it}^{*} > 0 \end{cases}$$

In this case a panel probit model is fitted to estimate the probability of holding status. We observe the purchase quantity of y_{it}^{*} in our data as well,

$$y_{it} = \begin{cases} 0, y_{it}^{*} \leq 0 \\ y_{it}, y_{it}^{*} > 0 \end{cases}$$

In this case a panel tobit model or truncated regression can be applied to explain how the quantity purchased varies across different characteristics. x_{it} is a set of exogenous variables which have controls such as age, gender, family status, job type, etc.

In equation (1a), h_{it}, which denotes self-reported health status, is the important variable of interest. Not only is this variable related closely to the motivation of purchasing supplementary health insurance, econometrically it is complicated by the endogeneity problem, which is a consequence of measurement error.

In GSOEP data, the respondents are asked to rate their own health on a 5-point scale where 1 means excellent, 2 good, 3 satisfactory, 4 poor, 5 bad. This categorical variable gives rise to the endogeneity problem as actual health is continuous instead of discrete. Therefore the discrete variable measures health with error.

Besides the fact that the discrete measure of health status gives rise to the measurement error, another concern with self-reported health is that respondents may have different reference points as to how they judge their health. Therefore there exists a gap between true health and the reported health where the reporting behaviour differs among respondents with different characteristics. For example, Crossley and Kennedy (2002) examine the reliability of self-assessed health using Australia data. They find that the measurement error is strongly related to the observable variables such as age, gender, income and etc[46]. Groot (2000) and Van Doorslaer & Gerdtham (2003) find that older respondents tend to have a "milder" view of their health and tend to rate their health as better than otherwise comparable younger respondents. In other words, the perception of what is "good health" varies across different age groups. This perception can be also varied between the rich and poor, employed and unemployed, optimistic and pessimistic individuals.

Let h_{it}^{*} denote the true health status and h_{it} the reported health status. The true relationship between these two can be written as:

$$h_{it} = h_{it}^{*} + \mu_{it}$$

Assuming μ_{it} is uncorrelated with the true health status, i.e, that μ_{it} is completely random, then we are facing a classic measurement error problem. Equation (1a) now becomes:

$$(1a') \quad y_{it}^{*} = \beta_1 x_{it} + \alpha * (h_{it}^{*} + \mu_{it}) + \varepsilon_{it}$$

[46] Individuals are asked to rate their health twice, the second time with an additional set of health related questions. They find there is about 28 percent change in their ratings and therefore it is reasonable to argue that self-assessed health is measured with error.

The measurement error in health status, μ_{it}, will result in an attenuated estimate of health effect on supplementary health insurance holdings.

To control the measurement error problem of the self-reported health status, using instrumental variables is the most commonly proposed method in the literature and econometric textbooks. However, finding the right instrumental variables is a hard task not only because they must satisfy the strict econometric conditions but also because they must be available in empirical data. Up to now, we find that most of the studies involving self-reported health status from GSOEP do not control for measurement error. Knaus and Nuscheler (2002) examine the effect of health on the probability of changing of health insurance with the GSOEP data from 1995 to 2000. They adopt a simultaneous two equation system in their empirical estimation. Self-reported health status is collapsed from 5 to 2 categories, good or bad, and then is estimated by a couple of control variables, such as doctor visits, hospital stay, long periods of absence from work due to illness, and other control variables are also included such as age, gender, and income. They use a Bivariate Probit model that jointly estimates health status and changes in health insurance. First, they estimate health status with ordinary probit and then the fitted values are used in a transition equation which is again estimated with ordinary probit. While the effects of health become significant with this approach, we suspect that the measures (doctor visits, hospital stay, long periods of absence from work due to illness) in their empirical model are not efficient in controlling for measurement error as these variables are more likely to be related to the decision to change health insurance. Furthermore, their empirical model which ignores the panel structure might yield incorrect estimates.

I find more objective measures of own health in GSOEP but they are only available in wave 2002 and 2004. The respondents were asked to describe their health status in daily activities such as whether they have trouble climbing stairs, whether they have difficulties dressing alone, whether it is difficult or they need help getting in and out of bed, whether they need help with shopping, whether they have difficulties doing housework alone. They were also asked to report their mental and physical health in last four weeks such as whether they are stressed out,

whether they are melancholy, whether they have some physical pain, whether they achieved less due to mental health or whether they are limited socially due to health. These variables could be used as instruments as they are correlated with self-rated health in such a way that they affect respondents' reference point when they report their health status during interview, but are uncorrelated with the error term in holding supplementary health insurance. A similar approach of modelling health status can be found in Bound et al. (1999). They use the detailed health information available from the Health and Retirement Survey to instrument for the endogenous and error-ridden self-reported health status. The predicted values are used as proxies in analyzing the labour market behaviour in the latter part of working life.

The empirical model I use to control for measurement error in self-reported health status consists of two equations in panel format:[47]

(1) $h_{it}^{*} = \beta_1 x_{1it} + \alpha_1 z_{it} + \varepsilon_{1it}$

(2) $y_{it}^{*} = \beta_2 x_{2it} + \alpha_2 h_{it}^{*} + \varepsilon_{2it}$

The first equation models the individual self-reported health status. Latent true health, denoted by h_{it}^{*}, is estimated by a set of objective health measures z_{it}, such as the daily activity indicators. x_{1it} is a set of exogenous variables including age, gender and income, etc.

The set of exogenous variables in Equation (2) contains some variables that do not appear in Equation (1). In equation (2), x_{2it} is a set of control variables such as the personal characteristics (age, sex, education), family background (married or single, number of children), social status (income, type of job, foreigner, a dummy for West Germany), insurance type (in AOK, or BKK, or other funds, whether the individual has compulsory insurance, voluntary insurance, family co-insurance, or insurance as a pensioner, unemployed person

[47] I do find (as shown in the empirical results section) the estimates are quite different between panel models and the models simply pooling observations.

or student), whether he has changed health insurance in previous year, his disability status, his willingness to take health risks, and his health care utilization in the previous year.

The nonlinear model controlling for measurement error given by Equations (1) and (2) in principle can be estimated by a two-stage maximum likelihood procedure. First, health status in Equation (1) is estimated by an Ordered Probit. Second, the fitted values of health status from the first step are used in the structural Equation (2). The estimation of the structural equation is again obtained by using a random effect panel probit model (when the dependent variable is a dummy) and a random effect panel tobit model (when the dependent variable is a quantity). Because I use predicted values of health status, the estimated standard errors in the second stage are incorrect. I apply a bootstrap technique, which takes into account the panel structure to correct standard errors.[48] Due to the complexity of the estimation technique, I ignore the potential correlation across time between the errors terms, ε_{it} and between measurement errors, μ_{it} .

In addition to the empirical models of probability and quantity purchased of supplementary health insurance, I model the supplementary contract that covers hospital stays. As the data is missing for those who do not have supplementary health insurance, the Heckman selection model is applied. In this case, the model of our interest is:

$$(3) \quad y_{2it}^{*} = \beta_3 x_{3it} + \alpha_3 h_{it}^{*} + \varepsilon_{3it}$$

where y_{2it}^{*} is the unobserved latent variable denoting the status of holding supplementary health insurance covering hospital stays. In our data we observe y_{2it}, which is a dummy that takes on a value of 1 or 0.

[48] My estimation strategy of using health proxies in the second step is similar to Bound et al. (1999). In addition, I use bootstraping to adjust the standard error in the second step.

$$y_{2it} = \begin{cases} 0, y_{2it}^* \leq 0 \\ 1, y_{2it}^* > 0 \end{cases}$$

I then add a selection equation to our model of interest, which is essentially the observed status of having supplementary health insurance:

$$(4) \quad y_{1it} = 1\left[\beta_2 x_{2it} + \alpha_2 h_{it}^* + \varepsilon_{2it} \geq 0\right]$$

when y_{1it}=1 we observe y_{2it}, otherwise we do not. My estimation strategy is to use the fitted values of health status in Equation (3) and apply the Heckman selection model. The standard errors are bootstrapped. The potential correlation between the errors terms across time is ignored. I perform the test of sample selection bias in Equation (3) as well.

IV. Data and Descriptive Statistics

Our data source is German Socio-Economic Panel Studies (GSOEP). GSOEP was started in 1984 as a longitudinal survey of private households and persons aged 18 years and older in the Federal Republic of Germany. It collects a rich array of information such as individual characteristics, social backgrounds, economic status, religions, personal opinions and attitudes toward some specific topics, etc. The original sample includes 4,528 households and all the full-age members in each household. In later years various refreshment samples were added. For example, in June 1990, 2,179 households from former German Democratic Republic (GDR) were included immediately after the reunion. In 1998, a refreshment sample of 1,067 households was added and in year 2000, 6,052 additional households were added.

As this micro data panel provides extensive information on the individual characteristics which are needed to analyze health and health insurance choice in Germany, my purpose is to explore this information to analyze the holdings of supplementary health insurance. The issues of interest include the characteristics of those who are more likely to have supplementary health insurance, whether the

story of adverse selection applies to the demand for supplementary health insurance in Germany, and whether there is risk selection in the provision of supplementary health insurance which results in market inefficiency.

The subjects studied in this chapter are all publicly insured individuals. As mentioned in the introduction, the majority of Germans are covered by the social health insurance. Since health care reforms in recent decades slashed more and more services and benefits provided by social health insurance, a growing proportion of the publicly insured members started looking for supplementary health insurance in the private health insurance market.

Figure 3.1 gives a picture of the supplementary health insurance trend from 1996 to 2005 in our data source. The percentage of supplementary health insurance holdings among the publicly insured has been increasing over this 10-year period. In 1996 it was about 4 percent and in 2005, it reached at around 13 percent. Over time the market for supplementary health insurance is expanding.

Figure 3.2 provides a more detailed picture about supplementary coverage.[49] Among the five categories of hospital stay, dental care, eye and corrective devices, travelling abroad, and others,[50] we can see that hospital stays are more frequently demanded than the rest. My reasoning is that hospital expenses are a large portion of household expenses once a family member has to get in-patient treatment, and therefore the household is more willing to insure against its financial threat.

Table 3.1 reports the summary statistics of variables used both in Equations (1) and (2) and the variables additionally used in Equations (1) and (2). I restrict the sample to the population between age 20 and 80, and I select waves 2002 and 2004 due to the availability of objective health-related information.

The average age is around 47 years. Fifty-three percent of the sample are females. The average education is 11 years and the mean household pre-tax income is 39590 Euro. I use household income instead of the individual income since the decision to purchase supplementary health insurance within the

[49] The period is from 1999 to 2005 as detailed coverage information was first available in 1999.
[50] This category includes the supplementary coverage for seeing medical practioners, which is not included in the usual health insurance contract.

household is more likely to depend on household financial ability. Around 8 percent of the sample is unemployed. Eleven percent reported having disability status. The average doctor visits in the previous year are around 10 times per year and the average hospital nights in previous year are around 1.66. Two percent reported that they incurred the out-of-pocket expenses for visiting doctors, therapists or non-medical practitioners in 2002. Around 3 percent employed reported that they were sick from work for more than six weeks at least once in the previous year. The average risk measure of an individual's willingness to take health risks is around 3.85 out of 11-point scale (where the willingness is increasing along the points). In general, the surveyed respondents are more risk averse with respect to health.

Among the additional variables used in the health status model are more objective health measures. I construct the individual's BMI out of their body weight and height and then classify them into three categories: underweight, normal and obese. Around 2 percent is underweight and 15 percent is obese in our data. Smoking is also included as it is a negative indicator of health status. The detailed physical and mental health information 4 weeks prior to the interview are categorical variables ranging from 1 (very often happened) to 5 (never happened). The particular health situation of a respondent might affect his reference point as how he rated his health status.

I examine in the empirical model three dependent variables of interest: a dummy for holding supplementary health insurance, the amount of supplementary health insurance (left truncated from 0 to a maximum of 800 Euro), and a dummy indicating hospital stay supplementary coverage. I include as explanatory variables insurance status that indicates whether an individual is a compulsory member (whose income is under a certain level) or a voluntary paying member, whether he has co-insurance as a family member or is insured as pensioner, whether he is unemployed, etc. I add controls for AOK and BKK as well. AOK (Allgemeine Ortskrankenkassen) plays a special role in the public health insurance system as it has to provide services to anyone who needs to be insured and not a member of other public sickness funds. Moreover, AOK mainly insures

blue collar workers who tend to bear a higher risk. BKK (Betriebskrankenkassen) is a company-based sickness fund and has, on average, lower contribution rates than all other funds. Consequently the adverse selection would result in more healthy insured in BKK.[51] I also include an indicator for whether the individual changed his health insurance provider in the previous year.

I define four age groups: the first group consists of individuals between ages 20 and 29, the second group consists of those between 30 and 56 (prime working age), the third group consists of those between 57 to 65 (pre-retirement age) and the fourth group consists of those older than 65.

Figure 3.3 presents the self-reported health status and the holdings of supplementary health insurance among different age groups. It is not surprising to see that self-reported health status diminishes among higher age groups. A higher proportion of the second and third group have supplementary health insurance than the other two groups. Despite its worse health status the group older than 65 has lower purchases of supplementary coverage because either the premium is too high to offset the benefits, or applications are rejected by the profit-maximizing private insurance companies.

Figure 3.4 presents a picture of the relationship between satisfaction with own health and the purchase of supplementary health insurance by the five categories of self-reported health status. Satisfaction with own health and the self-reported health status are no doubt positively related, as the connected line shows. The bar chart shows the percentage purchasing supplementary coverage by health status. It shows that healthier individuals are more likely to have supplementary health insurance. For example, 12 percent of the individuals who report excellent health have supplementary coverage, compared with 6 percent of the individuals who reported being bad health. While the theory of adverse selection suggests that low-risk consumers consume less insurance, that is not the phenomenon we observe here. One explanation I can offer is that private health insurance companies have access to applicant's data through hospitals and gatekeepers

[51] As concluded by Knaus and Nuscheler (2002).

(family doctors), which results in a market with less asymmetric information, mitigating the adverse selection problem.

Figure 3.5 shows the distribution of health status and the purchase of supplementary health insurance by 4 income quartiles. The first income quartile is the lowest 25 percent, and the fourth one is the highest 25 percent. From the left panel, we could see that individuals with higher family incomes were more like to report good health, while those with lower family income were more like to report bad health. There exists a positive relationship between income and health status, which could be interpreted as those with higher incomes are able to obtain more resources that can be allocated to health production and therefore have improved health.[52] This is in line with health production theory. The right panel shows a positive relationship between purchasing supplementary health insurance and income in general. The purchases within a quartile are stratified by different health status as well. For example, the first income quartile displays a pattern of higher levels of insurance purchase by healthier individuals. The second income quartile is more homogenous in holdings. The holdings of supplementary health insurance among the fourth income quartile is more concentrated among those who report a health status of excellent, good, satisfactory or poor. Combining both panels, the story from our data tell us that the poor are often in poorer health and do not have sufficient insurance coverage against health risks, which puts a high financial burden on them once the expensive health cost is incurred.

V. Empirical Findings

A. Baseline model

Table 3.2 shows results from the basic model (Eq.1a) for wave 2002 and 2004. I restrict the sample to individuals in both waves and therefore it is a balanced panel. I use the original self-reported health status to examine its effect on supplementary health insurance holdings. The random effect panel probit model controls for age, gender, education, marital status, number of children at

[52] The causality could be the other way around, i.e., the healthy are more productive in labour market and are able to earn higher income.

home, pre-tax household income, job status-white collar, blue collar, self-employed, or health related job, unemployed, self-reported health status, AOK or BKK member, insurance status, disability status, doctor visits, hospital nights and illness longer than 6 weeks last year, West German, foreigner, year dummy, and interaction between income quartiles and health status. In general, the holdings of supplementary health insurance are increasing for higher age groups but the increase is smaller in the oldest age group. Higher income is related to the higher holdings of supplementary health insurance.

The main variable I am concerned with is self-reported health status, which is negatively related to insurance holdings (better health is related to higher probability of holdings) but is statistically insignificant. This estimate is subject to bias due to the measurement error as discussed in the previous section. In the following sections, I instrument this variable with some additional information available in wave 2002 and 2004.

B. Health status equation

To start I calculate a Smith-Blundell test statistic for exogeneity after a simple probit model. [53] In the Smith-Blundell test, the null hypothesis states that the models are appropriately specified and all explanatory variables as exogenous. Under the alternative hypothesis, the suspected endogenous variable, self-reported health status, is expressed as a linear projection of a set of instruments, and the residuals from the first-stage regressions are added to the model. The residuals should have no explanatory power under the null and are included in the probit model. The test statistic is distributed Chi-squared (m), where m is the number of explanatory variables specified as endogenous in the model. I obtain a Smith-Blundell test statistic of 6.17, which is distributed as Chi-squared (1) with a P-value of one percent. Therefore I can reject the exogeneity of the self-reported health status and conclude that some work must be done to yield unbiased estimates of the effect of health.

[53] See Smith and Blundell (1986)

Table 3.3 reports the estimates of an Ordered Probit model of health status. I regress self-reported health status on a set of explanatory variables. First, the demographic variables appear to be significant in explaining health status. Higher age groups are associated with worse health status. Females are less likely to reported bad health. Married individuals are more likely to report bad health. Education year, however, appears to be statistically insignificant.[54] Higher income is positively related to better health outcomes, possibly due to the fact that more economic resources can be allocated to health production[55] (such as a healthier diet, sports equipment, better quality health care, etc.). Being unemployed is related to worse health, which confirms the detrimental effects of unemployment on health found in recent studies. Disability and being sick for longer than 6 weeks in the previous year are important indicators of health status. Health care utilization such as doctor visits and hospital nights in the previous year is positively related to worse health, as is out-of-pocket health expenditures. Being active in sports is defined as doing sports as least once a week. This variable turns out to be positively related to health outcomes. Smoking is positively related to poor health as expected. A higher body mass index (BMI) is associated with poor health, as indicated by the coefficient on obesity. Another set of variables which indicates certain physical or mental health conditions in the last four weeks prior to the interview are found to significantly influence an individual's reference point as to how they view their health status. For example, if the respondent was under stress in the last four weeks, he tended to report worse health status. Furthermore, the variables indicating whether an individual has trouble in daily tasks such as climbing up the stairs, dressing up, getting in/out of bed, shopping, doing housework alone, are positively related to worse health outcomes as well.

C. Supplementary Health Insurance Equation

I model the dependent variable-- holding of supplementary health insurance as a dummy - and estimate a random effect panel probit model with

[54] Using a dummy that indicates the level of education does not change the results.
[55] There can be reverse causality in the relationship as well if the healthy individuals are more productive in the labour market.

self-reported health status replaced by the fitted values from health status equation. The standard errors are bootstrapped. Table 3.4 provides the estimation results. The coefficients indicate positive or negative effects of the independent variables on the insurance probability.

The predicted health status turns out to be statistically significant at the one percent level and the effect becomes larger than when using the actual self-reported health status variable (-0.142 vs. -0.073). Measurement error in self-reported health status leads a downward bias. The estimation results show that the respondents in worse health status are less likely to have supplementary health insurance. It proves the existence of risk selection in the private insurance market for supplementary coverage. Through full access of a patient's record, the insurers select healthier individuals and leave the comparatively unhealthy out in order to optimize their risk structure. As the actual rejection rate is greater than zero, the equilibrium we observe is less than optimal.

The willingness to take health risks is positively and significantly related to supplementary health insurance holdings. However, it should not be interpreted as evidence of adverse selection because this variable is not equivalent to actual health status. In the health status equation, it turns out that the willingness to take health risks is insignificant in predicting health status. Furthermore, it is health status (not willingness to take health risks) which indicates the actual risk type to the insurers. While willingness to take health risks can capture the individual heterogeneity in the panel dataset, I also consider the possibility that the positive coefficient reflects reverse causality, i.e., the individual becomes more willing to take health risks after the purchase of supplementary health insurance.

Health care utilization such as doctor visits and out-of-pocket medical expenditures are positively related to the possibility of supplementary health insurance holdings. Other explanatory variable such as income is positively related to the holdings. Self-employed individuals are more likely to have supplementary health insurance, while the unemployed are less likely.

In the second step I estimate how the quantity of the supplementary health insurance responds to the different factors. In the dataset, the quantity purchased

ranges between 0 and 800 Euro per year. Essentially this variable is left censored at zero. I apply a panel random effect tobit model to Equation (2) in the two step estimation and bootstrap the standard errors. The results are shown in Table 3.5. Among the total 30594 observations in our dataset, 2940 are uncensored and 27654 are left-censored. The coefficients are not the marginal effects of the explanatory variables on amount of supplementary health insurance purchased, but instead the directions of their effects. In addition, I calculate in Table 3.5 the marginal effects on the conditional expectation and unconditional expectation, where the former is defined as

$$\frac{\partial E(y|y>0,x)}{\partial x_j},$$

and the latter is defined as

$$\frac{\partial E(y|x)}{\partial x_j}.$$

Conditional on supplementary health insurance purchase being positive, the person who is older than 65 is estimated to spend about 8 Euros more on supplemental health insurance. If we account for people who initially do not have supplementary health insurance, as well as those who have the positive purchase, the marginal effect of this age group dummy is about 12 Euros, which is larger than when we condition on a positive purchase. Income turns out to be positively related to the quantity purchased. The individual in the 4th income quartile on average spends around 13 Euros (unconditional marginal effect) more than their counterpart in the lowest income quartile. The self employed spend more on supplementary health insurance and the unemployed spend significantly less. The negative marginal effects of predicted health status tell us that people in poor health have less supplementary health insurance, which is consistent with the previous results. The willingness to take health related risks is positively related to

the quantity purchased. Again this variable underlines the individual's rational behaviour in seeking optimal protection with prior knowledge of his own risk type. Among the indicators of health care utilization, previous doctor visits turn out to be statistically significant and for every 10 visits are supposed to increase the conditional marginal purchase about 0.23 Euros and the unconditional marginal purchase about 0.5 Euros. The positive out-of-pocket medical expense, as well as whether the individual is sick for longer than 6 weeks in previous year increases the purchase as well. Germans, especially West Germans on average, purchase more supplementary health insurance than others.

Table 3.6 shows the estimates of the supplementary health insurance, which covers hospital stays using a Heckman Selection Probit model. As discussed earlier, hospital expenditures comprise not only the largest part of national health care costs but also impose a great impact on the financial status of an individual. The incentive to purchase additional coverage for hospital stays from private health insurance companies is greater than other kinds of coverage. I use a Heckman selection model because we only observe the status of having additional coverage of hospital stays among those who have supplementary health insurance. For those who do not have supplementary health insurance, this information is missing. If we set the value of the missing information to zero and estimate the whole sample, then we will underestimate the effect. If we restrict the sample to those who have supplementary health insurance, we will overestimate the effect. The test for selection bias in the regression yields a significant value of 17.04, which indicates that sample selection bias is present.

The evidence of risk selection is well illustrated by the negative coefficient on health status in Table 3.6. The marginal effects are calculated from the conditional probability, i.e., conditional on the positive purchase of supplementary health insurance. An integer-point increase in the predicted health status (health is worse) is associated with 3 percent decrease in the probability of having supplementary health insurance, which covers hospital stays.

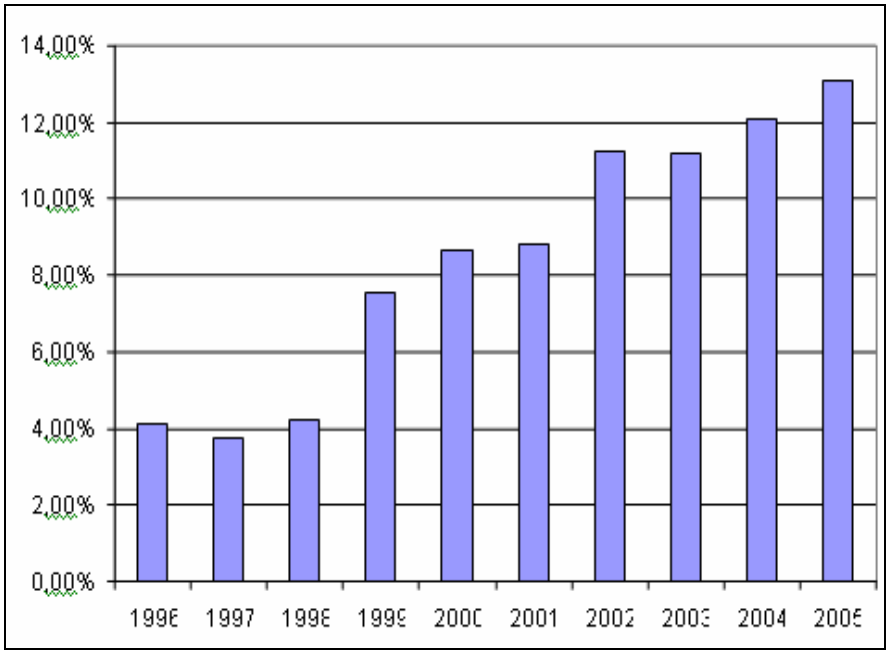
VI. Conclusion

This chapter examines the driving forces behind the holding of supplementary health insurance in Germany. Controlling for measurement errors, I mainly focus on the effect of health status on the status and quantity of holdings, as well as contracts covering hospital stays.

The results in this chapter find no evidence of adverse selection and instead confirm the existence of risk selection in private health insurance market for supplementary coverage. Not only those who are in poor financial status but also those who are in worse health status are more likely to be left without supplementary coverage. The current system which enables private health insurance companies to have full access to patient's data could explain the existence of risk selection. This is well illustrated from the empirical results that worse health status is negatively related to the holding of supplementary health insurance, especially with the coverage of hospital stays. The market provision of supplementary health insurance is not efficient in the sense that not all individuals who need supplementary coverage can purchase the contract in the private health insurance market.

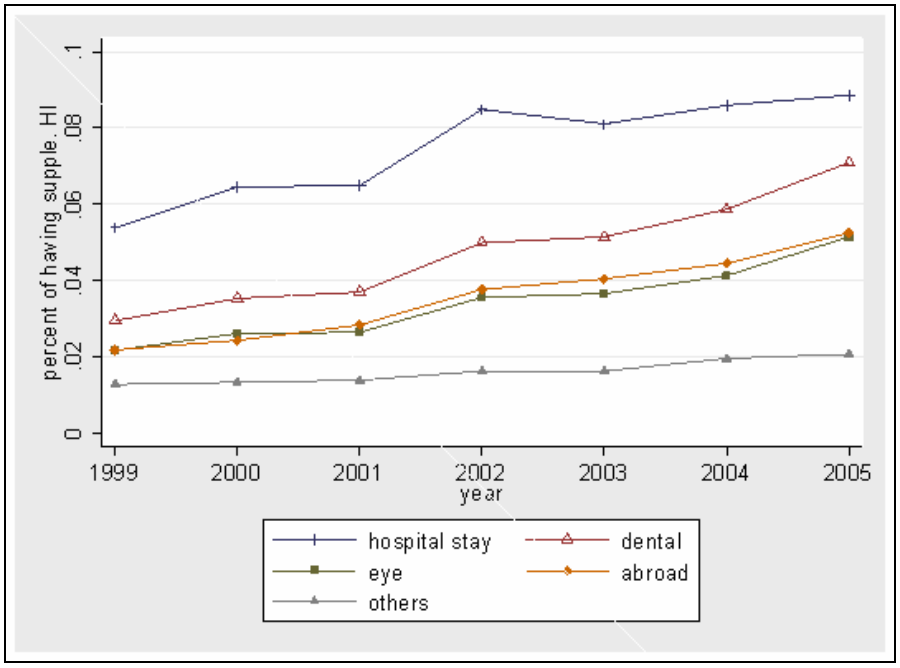
These results have implications on the current debate regarding patients' data protection. If risk selection prevents consumers from maximizing utility in the presence of health uncertainty, a specific policy favouring consumers' access to supplementary health insurance could improve market efficiency.

Figure 3.1: Purchase of Supple. Health Insurance among Compulsory Insured



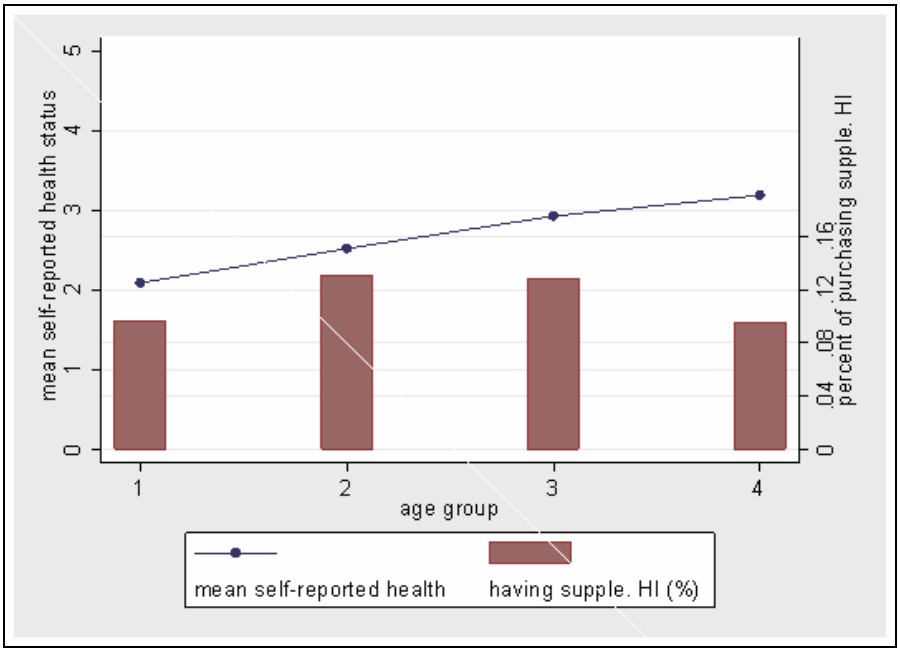
Source: German Socio-economic Panel Studies (1996-2005), author's self calculation.

Figure 3.2: Purchase of Supplementary Health Insurance With Different Coverage



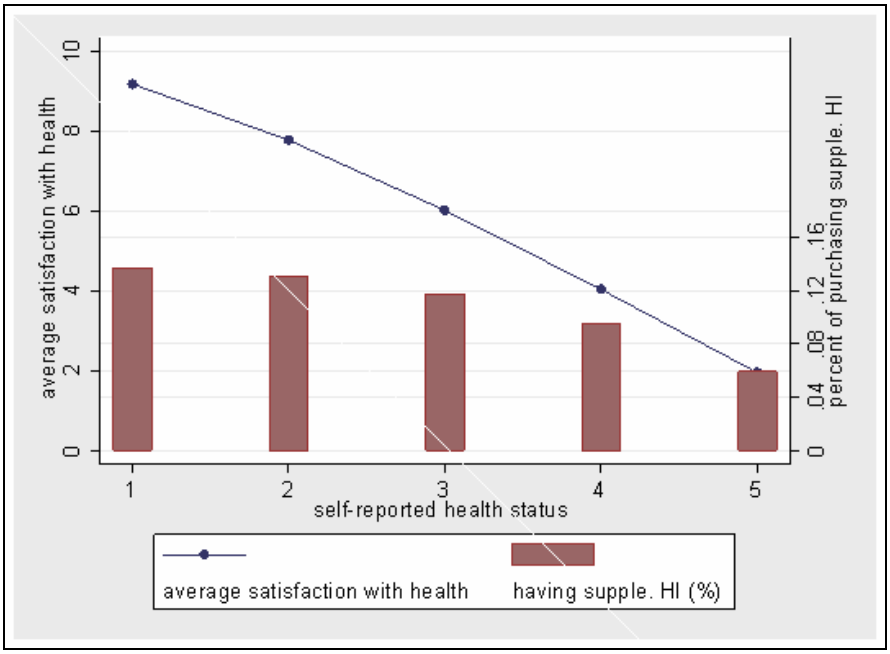
Source: German Socio-economic Panel Studies (1999-2005), author's self calculation.

Figure 3.3: Self-reported Health Status, Purchase of Supplementary Health Insurance among Different Age Groups



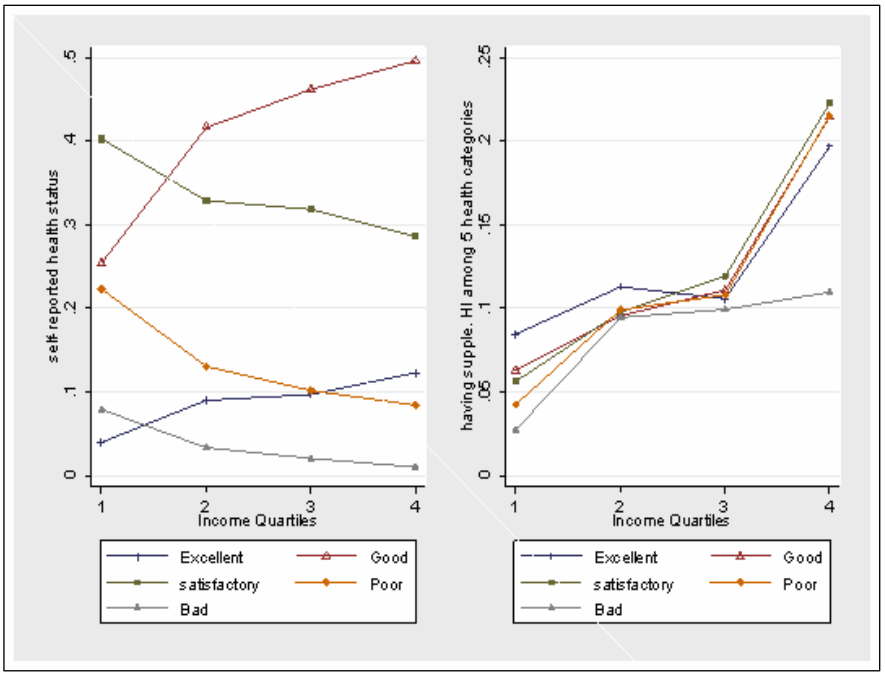
Source: German Socio-economic Panel Studies (wave 2002 and wave 2004)
Note: age group 1: age between 20 to 29
age group 2: age between 30 to 56
age group 3: age between 57 to 65
age group 4: age greater than 65

Figure 3.4: Satisfaction with Health, Purchase of Supplementary Health Insurance by 5 Categories of Health Status



Source: German Socio-economic Panel Studies (wave 2002 and wave 2004)
Note: self-reported health status --1: excellent,
--2: good,
--3: satisfactory
--4: poor
--5: bad

Figure 3.5: Distribution of Self-reported Health Status and Purchase of Supplementary Health Insurance by Income Quartile



Source: German Socio-economic Panel Studies (wave 2002 and wave 2004)
Note: income quartile -1: household pre-tax income first 25 percentile
-2: household pre-tax income between 25 and 50 percentile
-3: household pre-tax income between 50 and 75 percentile
-4: household pre-tax income above 75 percentile

Table 3.1: Summary Statistics of Variables

Variables used both in health status and supplementary HI model				
Variable	**Mean**	**Std. dev.**	**Min**	**Max**
Age	47.10	15.70	20	79
Female	0.53	0.50	0	1
Education	11.46	3.08	0	18
Married	0.66	0.47	0	1
Number of kids	0.56	0.92	0	9
Household pre-tax income[a]	39.59	41.39	0	1449.41
White-collar	0.32	0.46	0	1
Civil servant	0.01	0.08	0	1
Self-employed	0.04	0.20	0	1
Unemployed	0.08	0.27	0	1
Disability	0.11	0.31	0	1
Willing to take health risks	3.85	2.43	1	11
Doctor visits last year	9.88	16.27	0	360
# hospital nights last year	1.66	7.96	0	275
Medical expense in 2002	0.02	0.15	0	1
Sick longer than 6 weeks last year	0.03	0.18	0	1
Additional variables used in health status model Eq.(2)				
Active in sport	0.33	0.47	0	1
Underweight[b]	0.02	0.14	0	1
Obese[b]	0.15	0.36	0	1
Smoke	0.32	0.47	0	1
The followings occurred in last 4 weeks[c]				
Under stress	3.14	1.08	1	5
Melancholy, run-down	3.51	1.00	1	5
Well balanced	2.62	0.89	1	5
Use lots of energy	2.91	0.90	1	5
Physical pain	3.87	1.09	1	5
Achieved less due to physical health	3.87	1.09	1	5
Limited due to physical health	3.94	1.10	1	5
Achieved less due to mental health	4.15	0.99	1	5
Less thoroughly due to mental health	4.29	0.96	1	5

(to be continued)

Whether having trouble in				
Climbing stairs	0.40	0.49	0	1
Dressing	0.01	0.09	0	1
Getting in/out of bed	0.00	0.06	0	1
Shopping	0.01	0.11	0	1
Doing housework alone	0.01	0.10	0	1
Additional variables used in supplementary HI model Eq(3)				
Having supplementary HI	0.12	0.32	0	1
Amount of supplementary HI	5.08	24.11	0	800
Supplementary HI covers				
Hospital stay	0.09	0.28	0	1
Corrective device (e.g, glasses)	0.04	0.19	0	1
Dentures	0.05	0.23	0	1
Coverage abroad	0.04	0.20	0	1
Others	0.02	0.13	0	1
Insurance status				
Compulsory paying member	0.53	0.50	0	1
Voluntary paying member	0.12	0.32	0	1
Co-insured as family member	0.13	0.33	0	1
Insured as pensioner, unemployed, etc.	0.22	0.41	0	1
Health status	2.62	0.95	1	5
AOK member	0.33	0.47	0	1
BKK member	0.20	0.40	0	1
Changing health insurance last year	0.06	0.23	0	1
Year 2004	0.48	0.50	0	1
West German	0.745	0.44	0	1
German	0.92	0.29	0	1
Total observations	*30594*			

Data source: GSOEP 2002 and 2004 wave.

Note: a) in 1,000 euro.

b) underweight – BMI is less than 18.5; obese – BMI is greater than 30.

c) these categorical variables range from 1 (very often) to 5 (never).

Table 3.2: Panel Random Effect Probit Model of Supplementary Health Insurance Holdings

Supple. HI (dummy)	Coefficient	Prob>z
Age groups: based group -- age between 20 and 29		
Age between 30 to 56	0,292	0,006
Age between 57 to 65	0,495	0,000
Age greater than 65	0,621	0,000
Female	0,242	0,001
Education	0,102	0,000
Married	0,074	0,350
Number of Children	-0,142	0,000
Income quartiles: base group -- 1st income quartile		
2nd income quartile	0,704	0,000
3rd income quartile	0,908	0,000
4th income quartile	1,295	0,000
White-collar	0,036	0,680
Blue-collar	-0,313	0,002
Self-employed	0,429	0,003
Unemployed	-0,341	0,006
Self-reported Health Status	-0,073	0,173
Willing to take health risks	0,056	0,000
Disability	-0,092	0,397
Doctor visits last year	0,002	0,140
# hospital nights last year	0,000	0,919
Medical expense in 2002	0,301	0,134
Sick longer than 6 weeks last year	0,283	0,028
AOK member	-0,826	0,000
BKK member	-0,080	0,309

(to be continued)

Insurance status: base group--Compulsory paying member		
Co-insured as family member	-0,131	0,198
Insured as pensioner, unemployed, etc.	-0,380	0,000
Voluntary paying member	0,229	0,006
Changed HI provider last year	0,131	0,159
West German	1,133	0,000
German	1,291	0,000
Year 2004	0,184	0,000
Total Observations	*30594*	
Wald chi2(35)	*822.74*	
Log likelihood	*-8112.4985*	

Data source: GSOEP data Wave 2002 and 2004.

Table 3.3: Ordered Probit Model of Health Status

Health status	Coefficient	Prob>z
Age groups: based group -- age between 20 and 29		
Age between 30 to 56	0,383	0,000
Age between 57 to 65	0,588	0,000
Age greater than 65	0,663	0,000
Female	-0,092	0,000
Education	-0,003	0,859
Married	0,054	0,001
Income quartiles: base group -- 1st income quartile		
2nd income quartile	-0,071	0,001
3rd income quartile	-0,081	0,001
4th income quartile	-0,102	0,000
Unemployed	0,149	0,000
Willingness to take health risks	-0,001	0,816
Disability	0,318	0,000
Doctor visits last year	0,011	0,000
# hospital nights last year	0,010	0,000
Sick longer than 6 weeks last year	0,217	0,000
Medical expense in 2002	0,152	0,000
Active in sport	-0,146	0,000
Underweight	0,046	0,352
Obese	0,101	0,000
Smoke	0,089	0,000
The followings occurred in last 4 weeks		
Under stress	0,017	0,024
Melancholy, run-down	-0,109	0,000
Well balanced	0,077	0,000
Use lots of energy	0,265	0,000
Physical pain	-0,329	0,000
Achieved less due to physical health	-0,087	0,000
Limited due to physical health	-0,177	0,000
Achieved less due to mental health	0,025	0,013
Less thoroughly due to mental health	-0,042	0,000

(to be continued)

Whether having trouble in		
Climbing stairs	0,488	0,000
Dressing	0,278	0,043
Getting in/out of bed	0,489	0,008
Shopping	0,148	0,272
Doing housework alone	0,287	0,058
Total Observations	*30594*	
Log likelihood	*-28830,431*	

Data source: GSOEP data wave 2002 and wave 2004.

Note: other controls includes: four job type: blue collar, white collar, self-employed, health related job. Their coefficients are all statistically insignificant at the level of 10%.

Table 3.4: Panel Random Effect Probit Model of Supplementary Health Insurance Holdings

Supple. HI (dummy)	Coefficient	Bootstrapped Std. Errors
Age groups: based group -- age between 20 and 29		
Age between 30 to 56	0,348	0,107
Age between 57 to 65	0,586	0,132
Age greater than 65	0,735	0,158
Female	0,244	0,074
Education	0,098	0,013
Married	0,090	0,079
Number of Children	-0,139	0,039
Income quartiles: base group -- 1st income quartile		
2nd income quartile	0,694	0,115
3rd income quartile	0,893	0,127
4th income quartile	1,266	0,132
White-collar	0,029	0,085
Blue-collar	-0,313	0,103
Self-employed	0,418	0,145
Unemployed	-0,312	0,123
Health Status (Predicted)	-0,142	0,036
Willing to take health risks	0,057	0,014
Disability	0,006	0,113
Doctor visits last year	0,005	0,002
# hospital nights last year	0,001	0,004
Medical expense in 2002	0,326	0,205
Sick longer than 6 weeks last year	0,328	0,130
AOK member	-0,823	0,083
BKK member	-0,081	0,079

(to be continued)

Insurance status: base group--Compulsory paying member		
Co-insured as family member	-0,128	0,101
Insured as pensioner, unemployed, etc.	-0,380	0,089
Voluntary paying member	0,235	0,082
Changed HI provider last year	0,134	0,093
West German	1,127	0,088
German	1,294	0,165
Year 2004	0,185	0,035
Total Observations	*30594*	
Wald chi2(35)	*822.29*	
Log likelihood	*-8105.7231*	

Data source: GSOEP data wave 2002 and wave 2004.

Note: other controls includes: interaction between 4 income quartiles and health status (collapsed into good and bad two categories).

Table 3.5: Panel Random Effect Tobit Model of Amount Supplementary Health Insurance Holdings

Supple. HI (quantity)	Coefficient	Marg. effect on Conditional Expectation	Marg. effect on Unconditional Expectation
Age groups: based group -- age between 20 and 29			
Age between 30 to 56	18,157***	1,267***	2,751***
Age between 57 to 65	48,298***	4,595***	7,998***
Age greater than 65	70,062***	8,002***	12,394***
Female	10,037***	0,693***	1,502***
Education	3,986***	0,281***	0,606***
Married	4,531*	0,303*	0,662*
Number of Children	-4,947***	-0,338***	-0,730***
Income quartiles: base group -- 1st income quartile			
2nd income quartile	41,805***	3,997***	7,345***
3rd income quartile	45,805***	4,825***	8,625***
4th income quartile	68,144***	8,275***	13,308***
White-collar	-5,877*	-0,397**	-0,868*
Blue-collar	-20,213***	-1,283***	-3,001***
Self-employed	19,690***	1,658***	3,168***
Unemployed	-13,774***	-0,877***	-2,047***
Health Status (Predicted)	-5,618***	-0,425***	-0,919***
Willing to take health risks	2,007***	0,137***	0,296***
Disability	4,273	0,374	0,787
Doctor visits last year	0,318***	0,023***	0,049***
# hospital nights last year	-0,221	-0,014	-0,030
Medical expense in 2002	18,675***	1,506***	2,898***
Sick longer than 6 weeks last year	14,026**	1,030**	2,051**
AOK member	-33,464***	-2,126***	-4,903***

(to be continued)

BKK member	-6,728***	-0,446***	-0,989***
West German	52,381***	2,985***	7,434***
German	47,019***	2,315***	6,409***
Year 2004	4,678**	0,327**	0,705**
Insurance status--controlled			
Total Observations	*30594*		
Wald chi2(35)	*822.29*		
Log likelihood	*-8105.7231*		

Data source: GSOEP data wave 2002 and wave 2004.

Note: Marginal effect on Conditional Expectation is defined as $\partial E(y|y>0,x)/\partial x_j$

Marginal effect on Unconditional Expectation is defined as $\partial E(y|x)/\partial x_j$

Table 3.6: Heckman Selection Model of Supplementary Health Insurance Covering Hospital Stay

Supple. HI – Hospital (dummy)	Coefficient	Marginal effect
Age groups: based group -- age between 20 and 29		
Age between 30 to 56	0,158**	0,078**
Age between 57 to 65	0,551***	0,228***
Age greater than 65	0,600***	0,252***
Female	0,038	0,030
Health Status (Predicted)	-0,063**	-0,032***
Disability	-0,248***	-0,102***
Doctor visits last year	-0,001	0,000
# hospital nights last year	0,000	0,000
Sick longer than 6 weeks last year	0,033	0,032
AOK member	-0,114	-0,091***
BKK member	-0,210***	-0,094***
Changed HI provider last year	-0,133	-0,042
West German	0,540***	0,284***
German	0,034	0,081
Year 2004	-0,112***	-0,036***
Total Observations	*30594*	
Uncensored Obs.	*3553*	
Wald chi2(35)	*109.47*	
Log likelihood	*-11774.3*	
Wald test of rho=0	*17.04*	

Data source: GSOEP data wave 2002 and wave 2004.

Note: marginal effect is calculated based on the conditional probability -- Pr(supple. HI covering hospital stay | having supple. HI)

Reference

Angel, Ronald; Gronfein, William; *The Use of Subjective Information in Statistical Models. American Sociological Review*, Vol.53, No.3, pp.464-473, June 1988.

Atherly, Adam; *The Effect of Medicare Supplemental Insurance on Medicare Expenditures*, International Journal of Health Care Finance and Economics, Vol.2, No.2, pp.137-162, June 2002.

Belli, Paolo, *How Adverse Selection Affects the Health Insurance Market*, World Bank Policy Research Working Paper No. 2574, March 2001.

Bhandari, Shailesh; Health *Status, Health Insurance, and Health Services Utilization: 2001* Household Economic Studies, February 2006

Blomqvist, A., Johansson, P.-O., *Economic Efficiency and Mixed Public/Private Insurance*. Journal of Public Economics 66, 505-516, 1997

Bound et al. "*The Dynamic Effects of Health on the Labor Force Transition of Older Workers*". Labour Economics, 6 (1999) pp179-202.

Bound, John; Charles C. Brown and Nancy A. Mathiowetz. "*Measurement Error in Survey Data.*" PSC Research Report No. 00-450. August 2000.

Brown, Mark J. ; *Evidence of Adverse Selection in the Individual Health Insurance Market,* The Journal of Risk And Insurance, vol.59, No.1, pp.13-33, March 1992.

Brownstone, David; Valletta**,** Robert G.; **"***Modelling Earnings Measurement Error: A Multiple Imputation Approach*". The Review of Economics and Statistics, Vol. 78, No. 4 (Nov., 1996), pp. 705-717

Buonaccorsi, John; Demidenko, Eugene and Tosteson, Tor; "*Estimation Longitudinal Random Effects Models with Measurement Error*", Statistica Sinica 10 (2000), 885-903

Crossley, Thomas F. and Kennedy, Steven; *The Reliability of Self-Assessed Health Status*, Journal of Health Economics 21, 643-658, 2002

Hyslop, Dean R. and Imbens. Guido W. "*Bias From Classical And Other Forms Of Measurement Error*," Journal of Business and Economic Statistics, 2001, v19(4,Oct), 475-481.

Mullahy, J. 1997. "*Instrumental-Variable Estimation Of Count Data Models: Applications To Models Of Cigarette Smoking Behavior*," Review of Economics and Statistics, 79, 586-593.

Newhouse, Joseph P. et al, *The Demand For Supplementary Health Insurance, Or Do Deductibles Matter?* Journal of Political Economy, Vol.85, No.4, 1977

Nuscheler, Robert; Knaus, Thomas; *Risk Selection In The German Public Health Insurance System***.** Health Economics. Issue 14: 1253–1271, 2005.

Paolucci, Francesco; Prinsze, Femmeke; Stam, Pieter JA ; Wynand PMM van de Ven, *Solidarity in competitive markets for supplementary health insurance: an empirical analysis.* HEDG Working Paper 2006/02.

Pauly, Mark V. ; *Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection,* Quarterly Journal of Economics, 88(1) pp.44-62, February 1974.

Petretto, Alessandro; *Optimal Social Health Insurance with Supplementary PrivateIinsurance*, Journal of Health Economics, 18: 727–745, 1999.

Seidman, Laurence S.; *Supplementary Health Insurance and Cost-Consciousness Strategy*, The Journal of Risk and Insurance, Vol.45, No.2, pp291-310, Jun.1978

Showers, Vince E.; Shotick, Joyce A.; *The Effect Of Household Characteristics On Demand For Insurance: A Tobit Analysis,* The Journal of Risk and Insurance, Vol.61, No.3, pp492-502, Sep.1994.

Smith, Richard J.; Blundell, Richard W., *An Exogeneity Test For A Simultaneous Equation Tobit Model With An Application To Labor Supply*. Econometrica, 1986, 54:4, 679-686.

Windmeijer, F.A.G. and J.M.C. Silva. 1997. "*Endogeneity in Count Data Models: An Application to Demand for Health Care*," Journal of Applied Econometrics, 12, 281-94.

Zweifel and Breyer, *Health Economics*, Oxford University .Press, Oxford. 1997.

CURRICULUM VITAE

Personal Data

Name:	Joy Kapfer (formerly Jingyi Chen)
Date of Birth:	Oct.12, 1976
Place of Birth:	Fujian, China
Nationality:	German

Education

2004.7-2008.1	University of Munich, Germany, Munich Graduate School of Economics *Ph.D in Economics*
2002.8-2004.5	University of North Carolina at Chapel Hill, U.S.A. *Ph.D candidate in Economics*
2001.8-2002.7	East Carolina University, U.S.A *M.S in Economics*
1995.9-1999.6	Shanghai International Studies University, China. *B.A in Economics*
1992.9-1995.6	Wuping First High School, China
1989.9-1992.5	Wuping First Middle School, China

Research and Teaching Experience

2003.8-2004.6	University of North Carolina at Chapel Hill, U.S.A *Teaching Assistant*
2002.8-2003.5	University of North Carolina at Chapel Hill, U.S.A *Researching Assistant*
2002.1-2002.5	East Carolina University, U.S.A *Teaching Assistant*
2001.08-2001.12	East Carolina University, USA *Researching Assistant*