Master’s Thesis

Shuolin Shi

May, 2019

TITLE:	Variable Selection with Group Structure: Exiting employment at retirement age – A Competing Risks Quantile Regression Analysis*
AUTHOR:	Shuolin Shi (114133) MSc in Advanced Economics and Finance Copenhagen Business School
SUPERVISOR:	Ralf Andreas Wilke
SUBMISSION DATE:	May 15, 2019
NUMBER OF PAGES:	47
NUMBER OF CHARACTERS:	79,833

[*] We thank the Pension Research Centre at Copenhagen Business School for financial support and Kwang Woo Ahn for making sample code for the (adaptive) group bridge available.

Abstract

We consider the exit routes of older employees out of employment around retirement age. Our administrative data are high dimensional as they cover weekly information about the Danish population from 2004 to 2016 and 397 variables from 16 linked administrative registers, covering a wide range of information such as demographic, socioeconomic, financial, criminal, labor and health information, etc. We use a flexible dependent competing risks quantile regression model to identify how exits to retirement, illness, unemployment, etc. are related to the information in the various registers. To help finding an appropriate model we use variants of the Lasso, in particular the (adaptive) group bridge applied to competing risks quantile regression model to identify the relevant administrative registers and within-register variables. To our knowledge, this is the first application of these methods to large scale administrative data and the problem of exit into retirement. It is found that selected registers and most within-register variables from the (adaptive) group bridge have reasonable interpretation and remain significant in the unpenalized competing risks quantile regression. By applying state-of-the-art statistical methods to large scale data, we obtain detailed insights into the conditional distribution of transitions from employment into retirement in the presence of high dimensional data and competing risks setting.

Keywords: (Adaptive) group bridge; Dependent competing risks; Quantile regression; Retirement.

Acknowledgements

The past two years at CBS has been an invaluable experience for me. I grew and developed in so many ways that I could not imagine. This process would not have been possible without the help and support of many people and I feel deeply grateful to them.

First and foremost, my sincere thanks to my supervisor, Prof. Ralf A. Wilke, who has been giving me constant support, guidance and encouragement from the first semester of my study to the last part of this program. He always has insightful and detailed suggestions on the problem at hand and always teaches me with great patience. I have gained so much knowledge, experience and confidence regarding econometrics and statistical modelling from his courses and working with him over the past two years.

I would like to thank all the professors who have taught me in this program. They educated me with their great teaching skills, knowledge and personality. I was embraced in this creative, inspiring, equal and collaborative atmosphere and I benefit a lot from this.

I would like to thank my dear friends and classmates who keep pulling me out of the stressful postgraduate life.

I am grateful to my parents and grandparents for raising me up with love and care.

Last but not the least, my loving gratitude to my partner Xiaming, who ever stays by my side, loves me and supports me unconditionally. It is so lucky for me to be with her and so lucky for me to be who I am. I am forever grateful to her and grateful to all those experiences I had, all those people I met. They made me.

Content

Abstract i

Acknowledgements ii

1 Introduction 1

2 Literature Review 5

3 Institutional Settings 9

4 Data Description 12

4.1 Registers and Variables 12

4.2 Definition of Competing Risks 14

4.3 Computation of Duration 16

4.4 Sample Selection 17

5 Methodology 19

5.1 Variable Selection Theory 19

5.2 Competing Risks Quantile Regression 24

5.3 (Adaptive) Group Bridge in Competing Risks Quantile Regression . . . 27

6 Results 31

6.1 (Adaptive) Group Bridge 31

6.2 Competing Risks Quantile Regression 36

7 Conclusions and Discussion 43

References 45

A Appendix **49**

A.1 Tables 49

A.2 R Code 59

Chapter 1
Introduction

In an attempt to make the pension system fit for the future, politics in Denmark introduced more flexibility on the timing of retirement during the 2000s. This resulted for the employed in the possibility to decide on the retirement point once a certain age threshold is passed. This means, the point of retirement is no longer deterministic but possibly depending on a wealth of individual, economic and institutional factors. Previous analysis has mainly studied early retirement patterns, that is retirement before the official retirement age, while we consider both early and late retirement. The event to withdraw from employment may be in the discretion of the employed but can be also due to factors out of her control such as invalidity or dismissal. This calls for a model that permits for various exit routes. We adopt a flexible competing risks model that permits for dependencies between risks. By studying conditional quantiles, the determinants of an exit to the various routes are permitted to affect conditional distributions differently for long or short duration. This permits for important flexibilities as there are likely variables that play sizable roles at ages before the official retirement age but are not important at all afterwards and vice versa. In our analysis we combine state of the art distributional statistical methods with statistical regularization techniques to obtain a clearer picture about the factors that make people leaving their job earlier or later. It is common practice in social sciences and economics, that variables selection is undertaken sequentially or by means of variable inflation factors (VIF). However, these convenient approaches do not possess desirable statistical properties and therefore there are positive probabilities that selected variables do not belong to the model and that unselected variable actually belonged to the model even if the number of observations becomes very large. We adopt statistical approaches that permit for statistical regularization in high dimensional regressor spaces based on statistical learning. These techniques possess the oracle property and therefore have desirable statistical properties.

There is an extensive literature that considers transitions out of employment into (early) retirement (e.g. Lindeboom, 1998, Duval, 2003). Motivated by the ageing of the societies and subsequent restraints for the financial situation of the pension funds (Gruber and Wise, 1998), the question is analyzed how the institutional system can be shaped in order to avoid incentives to retire early. Beside direct (early) retirement, some form of other exit route through a bridging period in unemployment or disability may be chosen (see e.g. Miniaci and Stancanelli, 1998, Kyyra and Wilke, 2007, Fitzenberger and Wilke, 2010, Bingley et al., 2012). Relevant literature for Denmark has considered general determinants of retirement (e.g. Filges et al., 2012, Larsen and Pedersen, 2013, Kallestrup-Lamb et al., 2016). The use of duration models is limited to Christensen and Kallestrup-Lamb (2012), who study in particular the role of the health status. Existing studies for Denmark use annual data and most do not cover data for years after 2008. Therefore, they cover the period where a less flexible system was in place. Due to the low frequency only discrete time or discrete choice models have been applied.

We use linked administrative data provided by Statistics Denmark. The data contain weekly, monthly and annual observations for the period 2004-2016 of employees in Denmark. We select a subsample of employees aged 58 in 2008 who have stable working experience from 2004 to 2008 and are thus not assumed to retire due to limited career opportunities, limited work ability or long-term illness. There are 16 registers and 397 variables in total, forming 16 groups and 2 to 58 variables in each group. The registers contain information on various personal, household and firm characteristics, including demographics, education, income, pension, employment, socioeconomic status, health, criminal records and a wealth of company (employer) statistics. We consider two main exit routes, which are exits to retirement (via disability pension, early retirement pay and old age pension) and other exits (via unemployment, illness, death, etc.). 86% of the exits in our sample are to retirement. We compute employment duration at the age of 59 as the number of weeks until an exit takes place. The exit route is indicated by the status of the individual when employment ends and a gap of 4 weeks is allowed. If the individual enters into a retirement program or the other exit route before the end of employment, the first week of entrance rather than the last week of employment indicates the end of duration.

By using weekly information for a period of 8 years, we have (nearly) continuous duration data. This permits us the application of quantile regression. In particular, we study a dependent competing risks problem where an older employee can make a transition into retirement or another state. As the timing of duration is likely correlated through unobservables conditional on the observables, we adopt the dependent competing risks model of Peng and Fine (2009).

Since our data contain a wealth of linked registers and therefore numerous variables, it is not an easy task to select the relevant set of variables. In particular, many variables are categorical, such as industry information and geographic location. For this reason, we use variants of the Lasso (Tibshirani, 1996) to identify relevant variables in our model. The Lasso-type estimator is an attractive statistical approach because it has the oracle property, which means that the model selects only the relevant variables and the estimates for those variables equal the estimates from a model that only includes the relevant variables in probability under some regularity conditions. There is an extensive statistics literature about a number of variants of the Lasso that retain the oracle property in various settings. See Huang et al. (2012) for a survey. Zou (2006) introduces the adaptive lasso, which adds additional weights to the penalty to improve the selection. Huang et al. (2009) suggest the group bridge and Simon et al. (2013) suggest a sparse group lasso. These variants of the Lasso permit for group level and bi-level variable selection. Ahn and Kim (2018) combine the (adaptive) group bridge with competing risks quantile regression. We follow their approach in this paper. Being developed for problems and tested with data from medical sciences, it is unclear how it performs with our more heterogenous data structures. Our data are also characterized by a high degree of multicollinearities and it is to be seen how statistical learning can cope with this. The use of the Lasso in economic problems is still not widespread but increasing, though most applications are for standard mean regression or discrete choice models. Our quantile regression model is more complex as it is estimated separately for different quantiles. Therefore, the resulting set of variables changes by quantile and due to the upper bound of cumulative incidences, the model can only be estimated for a constrained set of quantiles. As the adaptive group bridge permits consistent identification of non-zero groups and within-group variables while maintaining the oracle property, we explore how adaptive group bridge is helpful in identifying relevant registers and within-group variables for the economic problem at hand. This provides guidance which registers are actually relevant for the problem under investigation.

Our study contributes to the literature as follows: It is the first study of these methods applied to large scale administrative data and analyzing exit into retirement. We are not aware that the (adaptive) group bridge has been applied in combination with competing risks quantile regression in the economics literature. For the analysis for Denmark we contribute by using weekly data, dependent competing risks and study a period where the latest retirement point is not deterministic. We explore the practical properties of combining a flexible and complex distributional model with statistical regularization methods.

The rest of the thesis is organized as follows. In chapter 2, we briefly review the existing literature about determinants of retirement in Denmark. In chapter 3, we give an overview of the main retirement programs in Denmark and some of their changes in order to extend working life in 2000s. In chapter 4, we describe the dataset including registers, individual variables, competing risks and durations, and then show the sample selection criteria and implementation issues. In chapter 5, we first briefly review existing techniques for variable selection, from sequential elimination to statistical regularization methods. Then we present the competing risks quantile regression framework and subsequently present the methodology we use in this thesis, which is a penalized competing risks quantile regression using (adaptive) group bridge following Ahn and Kim (2018). In chapter 6, we present and analyze the empirical results. In chapter 7, we give conclusions, discussions and suggestions for further research.

Chapter 2
Literature Review

In this chapter, we will review the literature about transitions to (early) retirement using Danish administrative data. Some studies focus on general determinants of retirement (e.g. Filges et al., 2012, Larsen and Pedersen, 2013, Kallestrup-Lamb et al., 2016), while others focus on a specific determinant (e.g. Danø et al., 2005, Christensen and Kallestrup-Lamb, 2012). Most studies focus on individual retirement, while some focus on joint retirement of married couples (e.g. An et al., 2004, Bingley and Lanot, 2007). Some studies focus on pathways to (early) retirement (Larsen and Pedersen, 2005, Bingley et al., 2012), while others focus on late retirement (e.g. Amilon and Nielsen, 2010) and semi-retirement (e.g. Larsen and Pedersen, 2013).

We find that for all these topics, most studies use discrete responses model (e.g. Filges et al., 2012, Larsen and Pedersen, 2013, Bingley et al., 2016, Kallestrup-Lamb et al., 2016) or failure time model (e.g. An et al., 2004, Christensen and Kallestrup-Lamb, 2012). Some studies only focus on aggregate statistics (e.g. Barslund, 2015, OECD, 2012a). Hardly any of the existing studies use data for years after 2008, when monthly employment statistics for employees (BFL) is available and a more flexible retirement system is in place, which could explain why most studies use discrete models with annual data. In short, to our knowledge, few studies have exploited the richness of comprehensive large scale individual administrative data and some studies appear to be relatively simple by today’s standards.

The uses of duration models include An et al. (2004) and Christensen and Kallestrup-Lamb (2012). An et al. (2004) study the joint retirement decisions of Danish married couples. Specifically, they examine whether the retirement timing of married couple is determined individually or jointly. The study is based on annual data for 243 working couples from 1980 to 1990. Despite the low frequency of the data, they specify a continuous time model by treating data as grouped. The multivariate mixed proportional hazard model that they introduce allows for both correlated unobserved heterogeneity and a positive probability on simultaneous termination of individual spells. In their model, each individual's retirement decision depends not only on their own characteristics but also on their spouses'. Each spell shares the same start time and is influenced by both common factors and individual factors. Results show that financial and health variables play a significant role in explaining both individual retirement and joint retirement; complementarities in leisure time explains joint early retirement decisions; correlation in unobserved heterogeneity, such as common tastes, plays a larger role than other observed heterogeneity in explaining joint late retirement decisions. Overall, retirement is a household decision.

Christensen and Kallestrup-Lamb (2012) study the determinants of duration until retirement, in particular the impact of changes in health status on early retirement behavior. The study is based on annual panel data for working people from 1985 to 2001. They use both single and competing risks specifications with both nonparametric and parametric baseline hazards in the grouped duration analysis. The model allows for time-varying regressors and flexible unobserved heterogeneity specification. They show that demographic, labor market status, financial variables and in particular health measured by objective medical diagnosis all have significant effects on retirement behavior. In the competing risks specification, they define five exit routes, including disability, early retirement, two kinds of unemployment and others out of the labor force. Results show that disability retirement and early retirement, unemployment followed by early retirement and by other programs differ significantly in terms of health and other regressors. Because in the single risk specification where all retirement programs are lumped together, opposite effects of a variable on different exit routes may cancel out and result in an insignificant estimator, competing risks specification leads to results that are more relevant in this study.

Another interesting study by Kallestrup-Lamb et al (2016) focuses on the general determinants of retirement using the adaptive Lasso applied to logistic regression. The study is based on annual data for working people for the year 1980 and 1998 and is the first application of Lasso-type estimators to this type and scale of data. They include 399 individual variables covering demographic, socioeconomic, financial, health and labor market status, the lags of time-varying regressors, and characteristics of the spouse if the individual is married. All types of retirement are lumped together. The penalized logistic regression model uses both the logit and Lasso estimator as initial estimator and possesses the oracle property. Results show that the choice of initial estimator for adaptive Lasso matters in terms of the number of selected variables; the effect of age, income, labor market indicators, wealth and health are stable over time, gender, marital status and different tuning parameter, suggesting that Lasso-type estimators give quite reasonable results.

Gørtz (2012) specifies a discrete-time proportional hazard model to study the early retirement behavior of female teachers in the day-care sector. In particular, she focuses on the role of working condition and health. The model uses a piece-wise constant baseline hazard duration framework and accommodates fixed effects to allow for unobserved heterogeneity. Results show that health and some measures of working condition have significant effects on early retirement decision for the period 1997-2006.

Several studies use option value model to calculate the potential gains of staying in the labor force. Danø et al. (2005) study the early retirement behavior of single women and single men respectively. They find that women are more willing to retire early than men and their retirement decisions are influenced by different variables: for men, income and health are the main factors, while for women, education and unemployment experience also matter. Bingley et al. (2004) study the impact of financial incentives on the probability of retirement. They find that the low-wage earners are incentivized to retire early and high-wage earners are incentivized to continue working. In the policy simulation study, they find that raising the eligibility age of retirement program will increase the average retirement age. Bingley and Lanot (2007) study the economic determinants of joint retirement behavior of married couples. They find that women are more influenced by changes in their own income and their spouses' income than men, and couples tend to retire early together due to complimented leisure. Bingley et al. (2016) study pension programs incentives by health and education level. Results show that economic incentives are generally important, and more important for people in poor health and low education level to retire early.

The remaining studies use linear probability model or discrete choice model. Gupta and Larsen (2007) study the effect of health shock on retirement for men and the role of welfare program in this effect. Results show that an acute health shock has a significant effect on retirement and almost none of the welfare programs could mitigate this effect. Gupta and Larsen (2010) further compare the effect of health on retirement between survey-based self-reported health and register-based diagnosis. They find that diagnosis is more important than economic factors; the retirement decisions for men and women are affected by different types of diagnosis; self-reported health yields biased estimator. Larsen and Pedersen (2005) aggregate three pathways from work to early retirement and analyze the probability of retiring through each pathway. They find that the determinants of retirement are different for each pathway: early retirement through the employment and unemployment are the dominant pathways and are affected more by availability of the program; other pathways are equally affected by individual characteristics. For semi-retirement and unretirement, Larsen and Pedersen (2013) find that demographic, education, unemployment experience, pension contributions and home ownership are significant factors and the effects are different for men and women. Filges et al. (2012) study the general determinants of retirement and focuses on the effect of unemployment in particular. They conclude that individual unemployment is highly significant and larger in magnitude than other demographic and education variables; program changes and cyclical situation also affect transition probabilities.

Chapter 3
Institutional Settings

Due to aging population, the public pension system is bearing higher financial pressure. In an attempt to make the pension system fit for the future, politics in Denmark with most European countries has introduced more flexibility on the timing of retirement during the 2000s in order to motivate working longer through, e.g. semi-retirement or late retirement, etc. In this chapter, we give an overview of the main retirement programs through public pension and labor market pension system and some of their changes in the 2000s in order to extend working life.

Old age pension (OAP), or state pension, is a universal pension that applies to every Danish national who has lived in Denmark for at least three years between the age of 15 and 65 and is aimed to protect the elderly from poverty. Besides the pension itself, pensioners are eligible for some other benefits, such as housing benefit, heating benefit, health-related benefit, etc. The retirement age has gone through several changes. For people born before 1 July 1939, the retirement age is 67; for people born between 1 July 1939 and 1 January 1963, the retirement age is shown in Table 3.1; for people born after 1 January 1963, the retirement age is according to future life expectancy.

The old age pension consists of a basic amount and a pension supplement and is means-tested. Although the test against income was reduced in the 2006 welfare reform, it is still unattractive to work and receive the state pension at the same time. For example, the pension supplement is reduced by 30% for annual income between 87,800 and 356,700 for singles and between 175,900 and 435,000 for individuals married/cohabiting with a pensioner. When the income is above the threshold, the pension supplement is canceled. From 1 July 2004, a pension deferral policy was introduced to motivate people to continue working after retirement age. People can postpone the state pension and get a higher payment afterwards if they work for at least 750 hours a year in the deferral period. Note that before 2011, the qualifying working hours was 1500 in 2004 and 1000 in 2008.

The maximum deferral period is ten years and people can defer the pension for two times.

Disability pension is another universal pension in Denmark, also called early Danish pension. It applies to Danish nationals who are between the ages of 18 and 65, have lived in Denmark for at least three years between the age of 15 and 65 and meet reduced work capacity criteria. After a disability pension reform in June 2012, the minimal age for disability pension is increased to 40 and people under 40 years old can only receive disability pension under special circumstances. Pensioners receive a certain amount depending on income level and receive some other housing and healthcare benefits. The reduced work capacity criteria require that due to social or health problems, the applicant is permanently unable to work under ordinary or flexible terms and unable to improve work capacity through treatment, activation, etc. An applicant needs to go through an assessment process and a rehabilitation program conducted by the municipal authorities to be entitled to the disability pension. The pension can be dormant or canceled, however, if in later periods the municipal authorities believe that the pensioner's work capacity is significantly improved.

Early retirement pension, or post employment wage (PEW) program, is a voluntary labor market pension. People can choose to be fully insured or partially insured. The scheme was introduced in 1979 in order to balance the unemployment of young people and the employment of older people. Pensioners have the opportunity to retire before the state retirement age and maintain a decent income level. Eligibility requires membership of an unemployment insurance fund, continuous contributions for at least 30 years, employment higher than 1,924 working hours or income higher than 233,375 within the last three years, and residence in Denmark. Similar to the state pension, the minimum retirement age for early retirement pension has gone through several changes as is shown in Table 3.1. For people born before 1 January 1954, the early retirement age is 60, which means that the maximum duration of early retirement is 5 years; for people born later, the retirement age is gradually increased to 65 and the duration is gradually reduced to 3 years.

The payment of early retirement pension differs according to previous income and insurance level, and is reduced if the pensioner has income from labor market pension, individual pension, work, etc. The maximum payment is the minimum of 90% of previous income and 91% of the unemployment insurance benefit rate. For people born before July 1 1959, they can choose to postpone the early retirement pay and get 100% of the unemployment insurance benefit rate, a tax-free premium for wages if they work for at least 1560 hours per year for fully insured and 1248 hours for partially insured, and start receiving early retirement pension no more than three years before the state retirement age. For people born before 1 January 1956, they can earn a set-off amount for other pension income in addition.

Other exit routes to early retirement include civil servants' pension, partial pension, etc. In this thesis we only consider the above mentioned three retirement programs as exit routes to retirement due to data availability.

Table 3.1: Retirement age of old age pension and early retirement pension

Date of birth	Old age pension	Early retirement pension
1 Jul 1939 – 31 Dec 1953	65	60
1 Jan 1954 – 30 Jun 1954	65.5	60.5
1 Jul 1954 – 31 Dec 1954	66	61
1 Jan 1955 – 30 Jun 1955	66.5	61.5
1 Jul 1955 – 31 Dec 1955	67	62
1 Jan 1956 – 30 Jun 1956	67	62.5
1 Jul 1956 – 31 Dec 1958	67	63
1 Jan 1959 – 30 Jun 1959	67	63.5
1 Jul 1959 – 31 Dec 1962	67	64
1 Jan 1963 –	68	65

Source: Borger.dk (2019)

Chapter 4
Data Description

The dataset we use is based on register data from Statistics Denmark (DST) and the DREAM database. It contains weekly, monthly and annual observations for 8178 employees born in 1949 for the period 2004-2016. We use 397 variables from 16 registers of DST's linked administrative data as regressors, and use DREAM to generate competing risks and durations. In this chapter, we first describe the registers and individual variables, then show how to define competing risks and compute durations. At last, we describe the sample selection criteria and implementation issues

4.1 Registers and Variables

Because individuals in the sample turn 60 in 2009 and start entering early retirement program, we use explanatory variables of 2008 to explain employment duration from 2009. We use 16 registers for the analysis. After converting categorical variables into dummy variables and going through some selection process, we have 2 to 58 variables for each register. Except for the employment statistics for employees (BFL) which has monthly information, the other registers are all observed annually. From the 16 registers, we have information covering a wide range of personal, household and firm characteristics, including demographics, education, income, pension, employment, socioeconomic status, health, criminal records and a wealth of company (employer) statistics, which could all possibly affect transitions from employment to retirement.

The 16 registers are the population statistics register (BEF, FAM), the education statistics register (UDDA), the criminal offences statistics register (KRIN), the health statistics registers (SGDP, SYIN), the income statistics registers (IND, LON), the pension statistics register (INPI), the labor market statistics register (AKM), the employment statistics for employees registers (BFL, IDAN, IDAP), and the company (employer) statistics registers (FIRM, FIDF, IDAS). Table A.1 in appendix A.1 contains descriptive statistics of the variables.

BEF is the population register. It contains information such as marital status, gender, geographical location, date of birth, citizenship, country of origin, household type, family type, etc. FAM contains number of children by age and number of people in the family. UDDA is the education register. It contains information on highest completed education and the institution. Education is divided into several categories following two classification rules: one is by subject area and the other is by level.

KRIN is the crime register. It contains the number, date and decision of criminal cases. Because there are only few cases each year and people are more likely to commit crime when they are young, we integrate the information from 1980 to 2008 to capture the long-run effect of criminal offenses. For the health registers, SGDP and SYIN, we integrate information from 2007 to 2008 for similar reasons. SGDP contains duration and amount of sickness benefits payments, duration of absence, and type of absence. SYIN includes diagnosis codes, treatment duration, geographical location, etc.

IND is the income register. It includes a large number of variables covering wage, tax, ATP contributions, debt, deposits, wealth, capital income, socioeconomic status, private pension, labor market pension, government transfer payments, etc. LON contains wage and employment statistics, including hourly wage, holiday pay, pension contributions, number of hours paid, number of holiday hours, number of paid absence hours, industry, occupation, sector, etc. INPI contains information on contributions to labor market pension and private pension schemes.

AKM describes the population's affiliation to the labor market throughout the year. It contains information on ATP contributions, occupation, working hours, industry and socioeconomic status. BFL contains detailed employment statistics based on SKAT’s eIncome register, including ATP contributions, wage income, number of hours paid, geographical location, sector, industry, etc., and is aggregated into annual regressors.

The IDA database links individual data with company data through the individuals' employment. IDAP contains individual employment information including working experience, number of (supplementary) jobs, working hours for the main/secondary job, insured level, insured duration, etc. IDAN contains information on employment duration, employment change, and employment type. IDAS contains workplace statistics including number of full-time equivalent employees, number of employees, industry, number of workplaces in the company, etc. FIDF contains company statistics including industry, sector, number of full-time equivalent employees, number of employees at the end of November, geographical location, ownership, etc.

Notice that some variables in different registers or within the same register contain the same or highly multicollinear values, which leads to high degree of multicollinearity. Our initial idea is to let (adaptive) group bridge select the relevant registers and variables. In the implementation stage, however, we find that high degree of multicollinearity is a problem for quantile regression even with Lasso. We choose to reduce multicollinearity beforehand by dropping variables with high VIF or with little variation. We also drop some variables with a large number of missing values, such as many company accounting statistics including gross profits, total assets, etc. Because this dataset only represents part of the information contained in each register, the relevant registers in the results should be interpreted with caution.

4.2 Definition of Competing Risks and Computation of Duration

The DREAM database is a comprehensive progress database based on the Ministry of Employment, the Ministry of Education, the CPR register and SKAT. The population of this database are individuals who received any government transfer payments from 1991. The type of payments is recorded weekly, including old age pension, disability pension, early retirement pension, unemployment benefits, sickness benefits, rehabilitation, cash benefits, wage subsidy, students’ grants, maternity benefits, etc. DREAM also contains employment information and many personal characteristics. Here we only use the government transfer payments records to identify retirement time, different competing risks and select the sample. The advantage of using DREAM is to have (nearly) continuous employment duration statistics.

We specify two competing risks: retirement and the others. The retirement exit route include all individuals who at least work until 4 weeks before they receive old age pension, disability pension or early retirement pension, which means that those who are semi-retired are also included and only the first entry is identified. Work here means having some positive working hours within a given month. The other exit route includes everyone else with exit to unemployment, illness, unknown labor market inactivity, death, etc. Table 4.1 reports the number and share of observed transitions into the two risks in the sample. We can see that 86% of the sample directly enter a retirement program from employment. Among those individuals, 3272 (40.01%) enter the old age pension, 3755 (45.92%) enter the early retirement pension, and only 16 (0.20%) enter the disability pension. The low occurrence of disability pension is quite reasonable due to the sample selection process shown in the next section. There exist only a small number of censored observations in the sample (end of data in 2016). This result, together with the result that most people enter old age pension within a few weeks after the state retirement age, is quite unexpected because quite a lot of evidence show that there are more and more people choose to defer pension and unretire (Amilon and Nielsen, 2010), and it further reinforces the idea that the sample in this analysis cannot represent the full population of employees.

Table 4.1: Number and share of transitions into risks

Risk	Number of observed spells	Share (%)
Retirement	7043	86.12
Others	944	11.54
Right-censored	191	2.34
Total	8178	100.00

4.3 Computation of Duration

We compute employment duration as the number of weeks from the first week of 2009 to the week when an exit takes place. The exit, or the end of employment, is identified by not working for two consecutive months. If the individual enters into a retirement program or the other exit route before the end of employment, the first week of entrance indicates the end of duration. If the individual enters into a retirement program or the other exit route after the end of employment, the last week of employment indicates the end of duration. Figure 4.1 shows the kernel density estimation of durations for right-censored observations and observations with exit to retirement. There is a sharp peak for durations between 260 and 315 weeks. That is because for this sample, those who enter old age pension do so within the first few weeks after they are eligible, at the age of 65, which corresponds to year 2014 and early 2015 respectively. Compared with old age pensioners, the distribution of duration for early retirement pensioners is more sparse. A smaller peak occurs during the second half of 2011, that is the time when the sample individuals reach the age of 62, suggesting that the early retirement pension deferral policy works well – people work until three years before the retirement age for better payment terms.

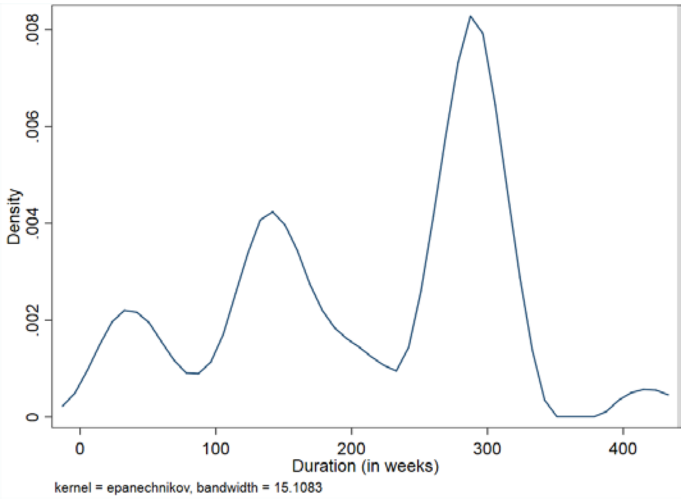


Figure 4.1: Kernel density estimation of durations[*]

[*] The underlying duration is a 5-person moving average due to policy concerning data confidentiality, yet this estimated density curve is very similar to the one without averaging.

On the contrary, few employees choose to defer old age pension, and those who choose to defer old age pension do so for a rather long period, at least until the end of 2016 when the data ends.

There is another way to compute durations. That is to follow the guidelines suggested by ILO (the International Labor Market Organization) – employment is prioritized higher than unemployment and other states outside the labor force, so people who work and receive retirement pension payments at the same time are categorized as working instead or retired. This will change the distribution of durations. The durations are more sparse over the whole period. 1344 more people are right-censored and there is a larger share of longer durations, which shows that many people choose to be semi-retired rather than unretired or completely retired. We do not use this specification because in this case, the censored observations will actually have two groups of people who keep working: one never enters a retirement program (unretired) and one enters a retirement program (semi-retired), which will bring some confusion.

Notice that there may exist measurement error when we need to convert the number of months from BFL into number of weeks to make it comparable with the weekly information in DREAM. The impact of this mismeasurement error is to be investigated.

4.4 Sample Selection

The idea of the sample selection process is to base the analysis on a group of employees who have stable work experience over a certain period and will not drop out of labor force due to limited career opportunities, limited work ability or long-term illness in the past. In this way, we avoid problems of modelling people who frequently transit between employment and unemployment and mitigate unobserved heterogeneity resulted from only using variables of year 2008 as regressors.

The stable work experience can be identified from the monthly working hour information in BFL. Since BFL only starts from year 2008, we also use the socioeconomic classification in AKM and government payments transfers in DREAM to supplement the selection process. The selection criteria include: (1) the individual is full-time employed every month for at least 11 months in 2008; (2) the individual is classified as employees every year from 2004 to 2008; (3) the individual does not receive disability pension from 1991 to 2008; (4) the individual does not enter the flex job scheme or service job scheme from 2004 to 2008; (5) the individual does not receive unemployment benefits or sickness benefits for more than 4 weeks in total every year from 2004 to 2008; (6) the individual does not receive unemployment benefits or sickness benefits consecutively for more than 4 weeks from 2004 to 2008.

Because the DREAM database only includes those who have received certain type of government transfer payments from 1991, it by definition does not include people who are censored – unretired throughout the observation period. In order to take into account of this selection bias problem, we use annual pension payments in IND and monthly working hours in BFL to identify censored observations. The selection criteria include: (1) the individual is full-time employed every month for at least 11 months in 2008; (2) the individual is classified as employees every year from 2004 to 2008; (3') the individual is not classified as pensioners from 2009 to 2016; (4') the individual does not receive retirement pension payments from 2009 to 2016. The impact of these different selection criteria is to be investigated.

Finally, we merge the regressors from register data with durations of the selected sample. It turns out that the population of each register is different and some variables within each register have missing values. For each variable with missing values, we could either drop this variable or drop those missing observations. Different choices lead to different datasets and we need to make many choices. The dataset of this analysis is chosen because it has a large number of observations and regressors at the same time. If the missing observations have patterns or characteristics that are correlated with the unobserved heterogeneity, we will have sample selection bias problem. This potential selection bias problem is to be investigated.

Chapter 5
Methodology

In this chapter, we introduce the methodology of this analysis. It is a combination of two parts. First, we briefly review the variable selection techniques to help understanding of the (adaptive) group bridge penalties. Then we introduce the competing risks quantile regression framework by Peng and Fine (2009). At last we present the methodology of this thesis – penalized competing risks quantile regression with (adaptive) group bridge by Ahn and Kim (2018).

5.1 Variable Selection Theory

Due to the development of computational power and availability of data, high dimensional regression has become more and more relevant for future research. There are two common situations in high dimensional data analysis where the classical maximum likelihood estimation fails: (1) The number of variables exceeds the number of observations; (2) There exists high degree of multicollinearity in the design matrix. In these situations, one might wish an oracle to reveal the relevant explanatory variables and use only those relevant variables for maximum likelihood estimation, and that is where variable selection methods play a role.

Traditional sequential elimination methods include subset selection based on likelihood ratio test or other tests and forward or backward stepwise selection based on information criteria such as Mallows's C_p, Akaike information criterion (AIC), Bayesian information criterion (BIC), etc. The problems with these methods include: (1) Only a small number of variables are allowed, because computation grows exponentially with a base of 2 as the dimension increases; (2) The results are unstable because the selection is discrete and different selection sequence leads to different selection results; (3) The resulted model may be overfitted and inaccurate, because post-selection inference based on maximum likelihood method ignores the error resulted from the selection process, and this is the main challenge for simultaneous variable selection and inference. In recent years, penalized regression and shrinkage methods have been introduced and gained much popularity when selecting relevant variables and carrying out statistical inference at the same time.

The objective function, or the minimization problem of the penalized regression is

$$Q(\beta|X, y) = L(\beta|X, y) + P_\lambda(\beta),$$

where $L(\beta|X, y)$ is the negative log-likelihood function which is the same as the standard maximum likelihood estimation, while for least squares estimation it is the sum of squared residuals, and $P_\lambda(\beta)$ is the additional penalty term that depends on the non-negative tuning parameter, or the regularization parameter λ, and the coefficients. The penalty term is the core of penalized regression methods. Using different penalty terms, we can assign different beliefs to the structure and magnitude of the variables and obtain different models in the end. The common belief for the following shrinkage and selection models is to penalize large coefficients.

Hoerl (1962) applies the ridge regularization method to regression analysis and the method is known as ridge regression afterwards. The penalty term of ridge regression is the ℓ_2-norm of coefficients,

$$P_\lambda(\beta) = \lambda||\beta||^2,$$

where $||\beta|| = \sqrt{\sum_{k=1}^{K} \beta_k^2}$, and K is the number of explanatory variables. To make the penalty unaffected by the scales of different regressors, all regressors need to be standardized beforehand. The ridge regression has a closed form unique solution and shrinks the coefficients towards zero. The regularization parameter controls the level of shrinkage. As λ approaches zero, the log-likelihood function dominates the objective function and the coefficients approach the maximum likelihood estimator. As λ approaches infinity, the penalty term dominates and the coefficients approach zero. The larger λ, the higher penalty and smaller coefficients. An advantage of shrinkage methods over traditional subset selection methods is that unlike subset selection methods where the change of model is discrete, the change in penalized regression models is continuous due to the continuity of regularization parameter. Many methods have been developed to choose the regularization parameter, including Mallows's C_p, cross validation, AIC and BIC. The criteria consist of two parts: One measures within sample fit and another measures the complexity, or the degrees of freedom of the model. By choosing a proper regularization parameter, the ridge regression estimator has smaller mean squared error (MSE) than OLS through a combination of a larger bias and a smaller variance. However, the ridge regression only has the shrinkage property – it cannot set any coefficient to zero and thus cannot be used to select relevant variables.

Tibshirani (1996) proposes the least absolute selection and shrinkage operator (Lasso). The penalty term of the Lasso is the ℓ_1-norm of coefficients,

$$P_\lambda(\beta) = \lambda\|\beta\|_1,$$

where $\|\beta\|_1 = \sum_{k=1}^{K}|\beta_k|$, and K is the number of explanatory variables. Using the absolute value of the coefficients as penalty terms, the Lasso can shrink all coefficients towards zero and set the coefficients of some variables to zero, thus achieving shrinkage and selection at the same time. The regularization parameter λ has similar properties as the ridge regression. The only difference is that as λ approaches infinity, the coefficients are all zero. Also, the Lasso does not have a closed form solution and the solution may be non-unique. Since the Lasso estimator is continuous in λ, we can draw a continuous coefficient path for all variables from the maximum value of λ where all coefficients are zero to the minimum value of λ. The methods to choose regularization parameter are the same as the ridge regression. Although under some conditions (Zhao and Yu, 2006), the Lasso selects relevant variables consistently, Fan and Li (2001) show that the Lasso has several shortcomings: (1) It leads to biased estimates and tends to overshrink large coefficients; (2) It tends to select irrelevant variables and thus has high false positive selection rates.

In order to reduce the bias and improve selection of the Lasso, Zou (2006) introduces the adaptive Lasso. The idea of adaptive Lasso is to assign different weights to the penalty terms such that large coefficients are penalized less heavily than small coefficients. The penalty term of adaptive Lasso is the same as the Lasso but with additional weights,

$$P_\lambda(\beta) = \lambda \sum_{k=1}^{K} w_k |\beta_k|,$$

where $w_k = 1/|\widetilde{\beta_k}|$, K is the number of explanatory variables, and $\widetilde{\beta_k}$ is a consistent initial estimator for β_k. If the initial estimator for β_k approaches zero, w_k approaches infinity and the adaptive Lasso estimator for β_k is zero. Zou (2006) proves that the adaptive Lasso with the maximum likelihood estimator as initial estimator has the oracle property under some regularity conditions. It is possible use different weight formulas and different initial estimators while retaining the oracle property. The (adaptive) Lasso estimator can select relevant individual variables, but does not perform well when variables have group structure, such as dummy variables formed from a categorical variable, or in our case, variables within each register. For variables with group structure, rather than identifying relevant individual variables, we sometimes only want to identify relevant groups and set the coefficients of all variables in the irrelevant groups to zero. But in this case, the (adaptive) Lasso also selects variables of irrelevant groups.

Yuan and Lin (2006) extend the Lasso to group variable selection and introduce the group Lasso. The idea is quite straightforward – to penalize on the group level instead of on the individual level. Suppose that the K explanatory variables can be divided into J groups. In each group there are A_j explanatory variables denoted by β_{jk} where $k = 1, \dots, A_j$. The penalty term of group Lasso is the ℓ_1-norm of groups with ℓ_2-norm of coefficients within each group,

$$P_\lambda(\beta) = \lambda \sum_{j=1}^{J} \sqrt{A_j} \left\| \beta_j \right\|,$$

where $\left\| \beta_j \right\| = \sqrt{\sum_{k=1}^{A_j} \beta_{jk}^2}$. The weights $\sqrt{A_j}$ adjust for sizes of groups and thus all else equal, groups with more variables will not be more likely to be selected than groups with less variables. The group Lasso reduces to the Lasso if all groups contain only one variable and thus shares the similar shortcomings of the Lasso on a group level. Because the ℓ_2-norm of coefficients is zero if and only if all the coefficients are zero, the group Lasso is able to drop an entire group. Within a group, the penalty term allows shrinkage, but does not allow selection on the individual level, similar to the ridge regression. So, the group Lasso either selects or drops all variables of a group. However, in many cases, only some variables within each group are relevant for the outcome and we want to include only those relevant individual variables in the analysis.

In order to select relevant individual variables with group structure, Huang et al. (2009) propose the group bridge method and further extend the Lasso to bi-level selection, which means to select both relevant groups and relevant individual variables within those groups. The penalty term of group bridge is a non-convex bridge penalty (Fu, 1998) applied to groups with ℓ_1-norm of coefficients within each group,

$$P_\lambda(\beta) = \lambda \sum_{j=1}^{J} A_j^{1-\gamma} \left\|\beta_j\right\|_1^\gamma,$$

where $\left\|\beta_j\right\|_1^\gamma = \left(\sum_{k=1}^{A_j} |\beta_{jk}|\right)^\gamma$, and γ is the bridge penalty that is between zero and one and set to be $1/2$ in their paper. By changing the penalty for within-group individual variables, an individual variable can be selected or omitted according to the effects from both itself and its group. The group bridge reduces to the Lasso if the bridge penalty is set to be one and there is only one variable in each group. Huang et al. (2009) prove the group selection consistency, but do not prove selection consistency for individual variables within groups. Due to the ℓ_1-type penalty, the group bridge shares similar shortcomings of the Lasso at individual level. There exist other types of non-convex penalties, such as SCAD (Fan and Li, 2001), MCP (Zhang, 2010), group MCP (Breheny and Huang, 2009), etc. See Huang et al. (2012) for a survey on group selection and bi-level selection methods.

It is still not an easy task to carry out inference with the Lasso-type estimator. Existing methods include sample splitting (Meinshausen et al., 2009), covariance test (Lockhart et al., 2014), post-selection inference (Lee et al., 2016), etc. See Taylor and Tibshirani (2015) for a survey. The inference method of this thesis is to use unpenalized regression models with the selected variables, which are believed to be the relevant variables if the Lasso-type estimator has the oracle property (Hastie et al., 2015).

There is an extensive statistics literature about a number of variants of the Lasso that achieve the oracle property in various settings, such as elasticity net (Zou and Hastie, 2003), sparse group Lasso (Simon et. al., 2013), fused Lasso (Tibshirani, 2005), hierarchical Lasso (Zhao et al., 2009), etc. See two books from Bühlmann and Van De Geer (2011) and Hastie et al. (2015) respectively for a comprehensive guide to statistical methods for high-dimensional data, with a focus on Lasso. The Lasso-type estimators have also been applied to various regression models, such as the standard linear regression, proportional hazards regression, logistic regression, etc. In this thesis, we use one variant of the Lasso that achieves selection consistency at both group level and within-group individual variable level – adaptive group bridge applied to competing risks quantile regression model, introduced by Ahn and Kim (2018), which is shown in chapter 5.3.

5.2 Competing Risks Quantile Regression

Koenker and Bassett (1978) introduce the quantile regression model. Instead of focusing on the conditional mean, they estimate the conditional quantiles of the explained variable. Quantile regression model is more complex than mean regression because it is estimated separately for different quantiles and gives a detailed analysis of the distribution of the explained variables. Let the conditional distribution of the explained variable Y be $F_Y(y|X) = Pr(Y \leq y|X)$ and X is a $N \times K$ matrix of regressors. The τ conditional quantile of $F_Y(y|X)$ is $Q_Y(\tau|X) = F_Y^{-1}(\tau|X) = inf\{y: F_Y(y|X) \geq \tau\}$. Assume a simple linear representation of the conditional quantile $Q_Y(\tau|X) = X\beta(\tau)$ where $\beta(\tau)$ is a vector of coefficients with length K. The estimator $\hat{\beta}(\tau)$ can be obtained by minimizing $\sum_{i=1}^{N}\left(\tau - \mathbb{1}_{\{y_i \leq x_i' b(\tau)\}}\right)(y_i - x_i' b(\tau)) \overset{\text{def}}{=} \sum_{i=1}^{N} \rho_\tau(y_i - x_i' b(\tau))$ with respect to $b(\tau)$, where $\rho_\tau(u) \overset{\text{def}}{=} (\tau - \mathbb{1}_{\{u \leq 0\}})u$ is known as the check function and $\mathbb{1}(\cdot)$ is an indicator function.

Quantile regression has been applied to various regression models, including linear, nonlinear, nonparametric model with cross section, time series and panel data. Duration data are usually censored and thus need special care. See Fitzenberger and Wilke (2015) for a survey on quantile regression methods and Fitzenberger and Wilke (2005) on quantile regression for duration analysis. Competing risks model refers to duration analysis with several potential failure types, or risks for one individual and only one failure type is observed if this observation is not censored. In this thesis there are two competing risks where an older employee can make a transition into retirement or another state.

Peng and Fine (2009) apply quantile regression to competing risks and introduce the competing risks quantile regression model. Different from independent competing risks proportional hazards model (Fine and Gray, 1999), this model can accommodate the dependency between competing risks, and thus allow duration to be correlated through unobservables conditional on the observables. Dlugosz and Wilke (2017) first apply this model to German maternity duration data and find that dependent competing risks quantile regression gives quite different results from independent competing risks proportional hazards model. We adopt this approach to have a flexible specification of dependencies between risks. With quantile regression, the regressors can affect conditional quantiles differently for long or short durations. This provides flexibilities further for heterogeneous effects on transitions for different durations or different retirement programs, which is exactly what we observe in chapter 6.2 that many variables play sizable roles in one’s early 60s but are not important at all afterwards.

Because we do not want to assume independent competing risks or specify a certain type of dependence form, the data generating process is unidentified (Peterson, 1976). The cumulative incidence curve avoids this problem by describing the distribution of observed transitions. Because it is not the marginal distribution of durations, the cumulative incidence curve is also known as subdistribution and cannot be higher than the share of observed transitions. In competing risks quantile regression, Peng and Fine (2009) use cumulative incidence curve to define the quantiles.

Consider a model with R types of competing risks $r = 1, \dots, R$. Let T_r and C denote event time and an independent censoring point, which in our case is the end of observation period and is a constant. Let $\epsilon = arg\, min_r\{T_r\}$ and $U = min_r\{T_r\}$. The observed duration is $T = min(U, C)$. The observed failure type is $\Delta = \mathbb{1}_{\{U \leq C\}}\epsilon$. The cumulative incidence for risk r is $F_r(t|X) = Pr(T_r \leq t, \Delta = r|X)$. The τ conditional quantile of $F_r(t|X)$ is $Q_r(\tau|X) = inf\{t: F_r(t|X) \geq \tau\}$. Assume $Q_r(\tau|X) = g\big(X\beta_r(\tau)\big)$, where $g(\cdot)$ is a known monotone link function and $0 < \tau_L \leq \tau \leq \tau_U < 1$. The sample analogue of (T, Δ, C, X) is denoted as $(t_i, \delta_i, c_i, x_i)$. In the case of no censoring, similar to the linear quantile regression model, the estimator $\hat{\beta}_r(\tau)$ can be obtained by minimizing $\sum_{i=1}^{N} \rho_\tau(g^{-1}(t_i^*) - x_i' b(\tau))$ with respect to $b(\tau)$, where x_i is a $K \times 1$ vector of regressors and $t_i^* = \mathbb{1}_{\{\delta_i = r\}} t_i + \mathbb{1}_{\{\delta_i \neq r\}} \times \infty$, which is equivalent to solving equation $N^{-\frac{1}{2}} \sum_{i=1}^{N} x_i' \left(\mathbb{1}_{\{g^{-1}(t_i) \leq x_i' b(\tau), \delta_i = r\}} - \tau\right) = 0$. In the case of independent censoring, $\hat{\beta}_r(\tau)$ is the solution to $S_N(b(\tau), \tau) = 0$, where

$$S_N(b(\tau), \tau) = N^{-\frac{1}{2}} \sum\nolimits_{i=1}^{N} x_i' \left(\frac{\mathbb{1}_{\{g^{-1}(t_i) \leq x_i' b(\tau), \delta_i = r\}}}{\hat{G}(t_i)} - \tau\right)$$

and $\hat{G}(t_i)$ is the Kaplan-Meier estimator for $Pr(C \geq T|X)$. Because $S_N(b(\tau), \tau) = 0$ may not have an exact solution due to noncontinuity, Peng and Fine (2009) define a generalized solution and show that the generalized solution is equivalent to minimizing the following ℓ_1-type convex function

$$\begin{aligned} U_N(b(\tau), \tau) = \textstyle\sum_{i=1}^{N} \mathbb{1}_{\{\delta_i = r\}} \left|\frac{g^{-1}(t_i) - x_i' b(\tau)}{\hat{G}(t_i)}\right| \\ + \left|M - b(\tau)' \textstyle\sum_{i=1}^{N} \frac{-x_i \mathbb{1}_{\{\delta_i = r\}}}{\hat{G}(t_i)}\right|, \\ + |M - b(\tau)' \textstyle\sum_{i=1}^{N} 2 x_i \tau| \end{aligned}$$

where M is a very large positive number. They prove consistency and asymptotic normality of $\hat{\beta}_r(\tau)$ under some regularity conditions. Besides, they propose consistent variance and covariance estimators, a trimmed mean statistic to summarize the effect over quantiles, and a constant test regarding whether a regressor has a constant effect on cumulative incidence quantiles. The trimmed mean effect estimator is represented as $\frac{\int_{\tau_L}^{\tau_U}\hat{\beta}(\tau)d\tau}{\tau_U-\tau_L}$ and a Wald-type test is derived for inference. In practice, we use Riemann sum to approximate the integral such that the estimated trimmed mean effect is $\frac{\sum_{\tau_L}^{\tau_U}\hat{\beta}(\tau)\Delta\tau}{\tau_U-\tau_L}$. For the constant test, the null hypothesis is $H_0{:}\,\beta(\tau)=\rho_0, \tau\in[\tau_L,\tau_U]$, where ρ_0 is an unspecified constant, and a test statistic is derived following the trimmed mean.

We use the R package *cmprskQR* by Dlugosz (2016) and revise one function for the estimation. See the revised function in appendix A.2. We use the exponential function as the link function $g(\cdot)$. The model is estimated for $\tau\in[\tau_L,\tau_U]$ with a step size of 0.01, where τ_L is 0.01 and τ_U is determined automatically as a value that corresponds to an cumulative incidence that is lower than its plateau value (Dlugosz and Wilke, 2017), resulting from condition C4 of Peng and Fine (2009).

5.3 (Adaptive) Group Bridge in Competing Risks Quantile Regression

Ahn and Kim (2018) introduce the adaptive group bridge and apply it to competing risks quantile regression. Similar to the change from the standard Lasso to the adaptive Lasso, the adaptive group bridge modifies the ℓ_1-type penalty of within-group coefficients to a weighted ℓ_1-type penalty. The setup is the same as group Lasso and group bridge mentioned in chapter 5.1. The penalty term of adaptive group bridge is

$$P_\lambda(\beta)=\lambda\sum_{j=1}^{J}A_j^{1-\gamma}\left\|\beta_j\right\|_1^{\gamma},$$

where $\left\|\beta_j\right\|_1^{\gamma}=\left(\sum_{k=1}^{A_j}w_{jk}|\beta_{jk}|\right)^{\gamma}$, $w_{jk}=\frac{1}{|\tilde{\beta}_{jk}|^{v}}$, $\tilde{\beta}_{jk}$ is an initial consistent estimator for β_{jk}, w_{jk} is the individual level weight for the k^{th} variable within group j and $v\geq 0$, $A_j^{1-\gamma}$ is the group level weight, and γ is the bridge penalty that is between zero and one. The group bridge is the case where $v=0$.

Combining the adaptive group bridge penalty with the objective function of competing risks quantile regression $U_N(b(\tau),\tau)$ in chapter 5.2, Ahn and Kim (2018) propose a penalized objective function

$$\begin{aligned} W_N(b(\tau),\tau) &= U_N(b(\tau),\tau) + P_\lambda\big(b(\tau)\big) \\ &= U_N(b(\tau),\tau) + \lambda \sum_{j=1}^{J} A_j^{1-\gamma}\left(\sum_{k=1}^{A_j}\left({|b_{jk}(\tau)|}\Big/{\left|\tilde{\beta}_{jk}(\tau)\right|^{v}}\right)\right)^{\gamma}. \end{aligned}$$

Minimization of $W_N(b(\tau),\tau)$ itself is not easy due to the non-convexity of this function. Similar to Huang et al. (2009), Ahn and Kim (2018) propose in Lemma 2.2 that through variable augmentation, minimizing $W_N(b(\tau),\tau)$ with respect to $b(\tau)$ is equivalent to minimizing $\widetilde{W}_N(b(\tau),\theta,\tau)$ with respect to $(b(\tau),\theta)$

$$\begin{aligned} \widetilde{W}_N(b(\tau),\theta,\tau) &= U_N(b(\tau),\tau) + \xi \sum_{j=1}^{J}\left(\left(\frac{\theta_j}{A_j}\right)^{1-\frac{1}{\gamma}} \sum_{k=1}^{A_j}\left({|b_{jk}(\tau)|}\Big/{\left|\tilde{\beta}_{jk}(\tau)\right|^{v}}\right)\right) + \xi \sum_{j=1}^{J}\theta_j, \\ \theta_j &= A_j^{1-\gamma}\left(\frac{1-\gamma}{\gamma}\right)^{\gamma}\left(\sum_{k=1}^{A_j}\left({|\beta_{jk}(\tau)|}\Big/{\left|\tilde{\beta}_{jk}(\tau)\right|^{v}}\right)\right)^{\gamma} \end{aligned}$$

where ξ is the tuning parameter and a reparameterization of λ. They prove that under some conditions, in the competing risks quantile regression framework, the group bridge selects group variables consistently; the adaptive group bridge not only selects group variables consistently, but also selects within-group individual variables consistently, and thus possesses the oracle property.

In the simulation study and real data analysis, Ahn and Kim (2018) set γ to be $1/2$ and v to be 1. Following them, we use these values in the algorithm of solving the minimization problem. The (adaptive) group bridge algorithm is:

1. Choose a certain quantile.
2. Use the group bridge estimator or the unpenalized competing risks quantile regression estimator as the initial estimator $\tilde{\beta}_{jk}(\tau)$ to compute the individual weights for adaptive group bridge estimator. For the group bridge estimator, the individual weights do not appear.
3. Following Friedman et al. (2010), choose a grid of 100 values for the tuning parameter

ξ_n that is uniformly spaced on the log scale. The upper bound is the smallest value where none of the variables is selected and the lower bound is the upper bound over 1000. For each value of the tuning parameter, repeat the following steps for $t = 1, \ldots$ until practical convergence indicated by $\left\|\hat{\beta}^{t}(\tau) - \hat{\beta}^{t-1}(\tau)\right\|_1 < 0.001$, and save the estimated coefficients $\hat{\beta}(\tau)$ after practical convergence:

a) Compute $\theta_j^{(t)} = \sqrt{A_j \sum_{k=1}^{A_j} \left(\left|\beta_{jk}^{(t-1)}(\tau)\right| \Big/ \left|\tilde{\beta}_{jk}(\tau)\right| \right)}$ for all groups $j = 1, \ldots, J$, where for the first iteration $\beta_{jk}^{(0)}(\tau) = \tilde{\beta}_{jk}(\tau)$.

b) Solve the minimization problem of (adaptive) group bridge

$$\begin{aligned}\hat{\beta}^{t}(\tau) &= \underset{b(\tau)}{argmin}\, U_N(b(\tau), \tau) + \xi_n \sum_{j=1}^{J} \left(\frac{A_j}{\theta_j^{(t)}} \sum_{k=1}^{A_j} \left(\left|b_{jk}(\tau)\right| \Big/ \left|\tilde{\beta}_{jk}(\tau)\right| \right) \right) \\ &= \underset{b(\tau)}{argmin}\, U_N(b(\tau), \tau) + \xi_n \sum_{j=1}^{J} \sum_{k=1}^{A_j} w_{jk} \left|b_{jk}(\tau)\right|\end{aligned},$$

where $w_{jk} = \frac{A_j}{\theta_j^{(t)} |\tilde{\beta}_{jk}(\tau)|}$.

4. Now we have 100 estimates $\hat{\beta}(\tau)$ for 100 values of the tuning parameter respectively. To choose the optimal tuning parameter, we compute the BIC-type criterion following Ahn and Kim (2018)*

$$\frac{2}{n} U_N(\hat{\beta}(\tau), \tau) + p_n\, ln(K) \frac{ln(N)}{2N},$$

where K is the number of explanatory variables, N is the number of observations, and p_n is the number of nonzero coefficients or selected variables to model degrees of freedom. The tuning parameter that leads to the smallest criterion value gives the optimal estimates $\hat{\beta}(\tau)$.

Notice that due to the combination of quantile regression and the variable selection technique, both the estimates and the selected variable set change by quantile. Due to the computational intensity, we only choose three different quantiles for estimation. After the

* Other methods exist. For example, Ahn et al. (2018) apply the generalized cross validation method following Huang et al. (2014) to choose the tuning parameter.

selection process, we use the unpenalized competing risks quantile regression for estimation and inference. Following Ahn and Kim (2018) and Peng and Fine (2009), we explore how (adaptive) group bridge is helpful in identifying relevant registers and within-register variables in competing risks quantile regression to obtain a complete view of the conditional distribution of observed employment durations with exit to retirement. R code for the (adaptive) group bridge estimator is provided in appendix A.2.

Chapter 6
Results

In this chapter, we present and discuss the empirical results. First we look at the estimation results from (adaptive) group bridge for three quantiles and make inference using unpenalized competing risks quantile regression with the selected variables. Then we present estimated coefficients and estimated cumulative incidence quantiles over equally-spaced quantiles to have a detailed view of the conditional distribution of coefficients and observed transitions into retirement respectively across quantiles.

6.1 (Adaptive) Group Bridge

For the (adaptive) group bridge estimation, we analyze three quantiles: 0.11, 0.25 and 0.31. For each of these quantiles, we use three selection methods: group bridge, adaptive group bridge using group bridge as the initial estimator, and adaptive group bridge using competing risks quantile regression as the initial estimator. Table 6.1 shows the estimation results. Only variables that are selected by at least one of the three methods are included in the table. The last column for each quantile is the estimation results from the unpenalized competing risks quantile regression. For better inference, we keep only one of the highly correlated variables, such as 'employed: highest level' and 'occupation: highest level'. We provide definition and links from Statistics Denmark of some variables in Table A.2 in appendix A.1.

We can see that for $\tau = 0.11$, all methods select 3 registers and 10 to 15 within-register variables. The group bridge estimator is more likely to shrink the magnitude of the estimates compared with adaptive group bridge; however, this phenomenon is less apparent for the other quantiles. The selection results for $\tau = 0.25$ is similar to those for $\tau = 0.11$, yet for $\tau = 0.31$, the results change considerably, only 1 to 3 registers and 1 to

4 within-register variables are selected. For all quantiles, adaptive group bridge with group bridge as initial estimator has the largest model size and group bridge estimator select the least. Estimates using (adaptive) group bridge all have the same sign for all quantiles. Most selected variables are significant in the competing risks quantile regression, but few have different signs. In all, all three methods reduce dimension considerably; although the selection results change across quantiles, the selected registers and within-register variables share some similarity; there is no clear comparison of selection quality among the three methods; the selection results have reasonable interpretation and are consistent with many studies.

Of the 16 registers, the education, health, crime and firm registers are not selected. We are left with registers covering labor market, employment, population and financial statistics. This seems to contradict some studies. But there are some explanations. First, education is known to have ambiguous effects on retirement. Second, the sample of this study consists of rather healthy individuals and from the data we know that only few people have criminal records, which could explain the omission of health and crime information. And there are actually few studies mention the effects of firm characteristics.

For the labor market register AKM, some occupation and industry variables are selected. Except people in the energy supply industry and in work that requires the highest-level skills, the other selected occupations and industries all have a negative sign, suggesting higher probability of transitions into retirement. However, it seems that most industries and occupations are not important for retirement transitions. Work experience, unemployment experience, etc. are also not selected possibly because the sample consists of employees with rather similar work experience and little unemployment experience.

For the population register BEF, the variable ‘date of birth’ is computed as the number of weeks from the first week of a year to the date when one is born in order to capture the effect of age eligibility. For example, the value of this variables is 5 if one is born on 2 February. It is selected by almost all methods and all quantiles and is highly significant, suggesting that the employment duration is very sensitive to the date a person is born. The positive sign shows that people born earlier has shorter duration than those who born later in a year, suggesting that many people retire just after they satisfy age requirement.

The variable ‘male’ is only selected by group bridge, and it appears insignificant in all quantiles. This finding contradicts many studies which show that gender plays an important role in retirement decisions. One reason could be that the effect of gender is wiped out by other variables, such as occupation, industry, etc. In particular the variable ‘reference person in family’, which refers to the women in a family of heterosexual couple and the oldest person in other families, by definition is highly correlated with gender. It is selected for the first two quantiles and is significant for quantile 0.11. The negative coefficient suggests that everything else equal, females or the oldest person in a family have a higher probability of observing a transition into retirement at shorter durations.

Adaptive group bridge with group bridge as initial estimator is the only method that includes IDAP register. There is only one variable ‘insured’ selected in this register, but it is highly significant in the competing risks quantile regression. This variable suggests whether one is insured or not, but it has a confusing description in Statistics Denmark, which makes the definition unclear. It is interesting that these methods select a variable that does not have a clear interpretation but seems to be important and survive the tests.

Adaptive group bridge with group bridge as initial estimator does not select the income register IND, and the other methods select within-group variables less consistently compared with AKM and BEF. Some variables are highly correlated, such as the variable ‘AM-income’ and ‘salary income’, etc., and thus is eliminated from the unpenalized regression. The income variables all have positive sign for all quantiles, suggesting that people with higher income tend to work longer, corresponding to the case where substitution effect dominates in the tradeoff between leisure and income. On the other hand, the variable ‘property value’ has a negative sign in the competing risks quantile regression model, showing evidence of income effect, consistent with Kallestrup-Lamb et al. (2016). Contributions to pension schemes, the PEW in particular, have a negative effect on employment durations for lower quantiles. Debt value is only selected by group

Table 6.1: Estimation results for employment durations with exit to retirement

Regi-sters	Within-register variables	0.11 Quantile				0.25 Quantile				0.31 Quantile			
		GB	AGB-GB	AGB-CRQR	CRQR	GB	AGB-GB	AGB-CRQR	CRQR	GB	AGB-GB	AGB-CRQR	CRQR
AKM	Occup.: Operation/Transport	-0.087	-0.298		-0.368‡		-0.183		-0.162				
	Occupation: Manual	-0.338	-0.730	-0.562	-0.548‡	-0.178	-0.571	-0.331	-0.055				
	Occupation: Military	-0.661	-1.261		-1.098‡	-0.485	-1.163		-1.276‡				
	Occupation: Highest level	0.308		0.277		0.233		0.375					
	Municipal employment	-0.302	-0.318		-0.215†	-0.204	-0.293	-0.129	-0.025*				
	Industry: Energy supply		0.227		0.377‡		0.281		0.082				
	Industry: Education			-0.110	-0.231*			-0.068	-0.034*				
	Industry: Healthcare			-0.292	-0.120			-0.126	-0.031*				
	Employed: Highest level		0.144		0.242‡		0.327		0.049‡				
	Employed: Basic level	-0.301	-0.511	-0.496	-0.286‡	-0.181	-0.365	-0.316	-0.005				
BEF	Date of birth	0.006		0.008	0.013‡	0.009		0.011	0.006‡	0.036	0.025	0.031	0.006‡
	Divorced	0.032	0.329		0.297‡		0.326		0.083‡				
	Reference person in family		-0.383	-0.064	-0.327‡		-0.365	-0.081	-0.029				
	Household: Married couple	-0.016			-0.178‡								
	Male	0.159			0.014	0.119			0.010				
IDAP	Insured		-0.834		0.329‡		-0.914		0.126‡				
IND	AM-income (million)	0.675		0.865	0.794‡	0.711		0.970	0.432‡	0.932			0.722‡
	Capital income (million)							0.581					
	Other capital inc. (million)	0.680			1.529†	0.538			-0.002				
	ATP contributions (thous.)			0.099	-0.042			0.033	-0.023‡				
	Contrib. to PEW (thousand)	-0.155		-0.185	-0.180‡	-0.197		-0.187	-0.132‡	-0.168			-0.128‡

(Continued)

Table 6.1: Estimation results for employment durations with exit to retirement (Continued)

Regi-sters	Within-register variables	0.11 Quantile				0.25 Quantile				0.31 Quantile			
		GB	AGB-GB	AGB-CRQR	CRQR	GB	AGB-GB	AGB-CRQR	CRQR	GB	AGB-GB	AGB-CRQR	CRQR
IND	Contributions to union, UI, and PEW (thousand)	-0.012			-0.017	-0.015			-0.007‡				
	Property value (million)			0.009	-0.042‡								
	Salary income (million)											0.925	
	Debt value (million)	0.023			0.130‡	0.033			0.074‡	0.059			0.092‡
	Taxable personal inc (mil.)					0.059							
	Pension income (million)			0.677	1.216			0.701	0.659†				
	Interest expense (million)					0.273							

Note: The values are estimated coefficients for from three different quantiles using five related methods. GB refers to the group bridge estimator. AGB-GB refers to the adaptive group bridge estimator using the group bridge as initial estimator. AGB-CRQR refers to the adaptive group bridge estimator using competing risks quantile regression as initial estimator. CRQR refers to the unpenalized competing risks quantile regression estimator. Only variables that are selected by at least one of the group bridge and adaptive group bridge methods are included in the table. Significance is indicated as: 10% (*), 5% (†), 1% (‡).

bridge but remain significant for all quantiles. The positive sign reflects the need to work longer for people who have more debt, which is reasonable if wage income exceeds retirement pension income.

However, we have to mention that the choice of tuning parameter plays an important role in the estimation process. If we change the tuning parameter, the selection results change. See Figure 6.1 for the value of BIC-type criterion for the adaptive group bridge estimator using competing risks quantile regression as the initial estimator for the 0.11 quantile. We can see that although there exists a global minimum, the curve is relatively flat for a wide range of λ around the global minimum, which means that we could get very different selection results if we use another tuning parameter which has very similar criterion value. The curves for other methods and other quantiles exist similar pattern. This pattern casts doubt on the optimal value for tuning parameter and the resulted selection results, so we should treat results in Table 6.1 with caution if we want to use the selected variables for further analysis.

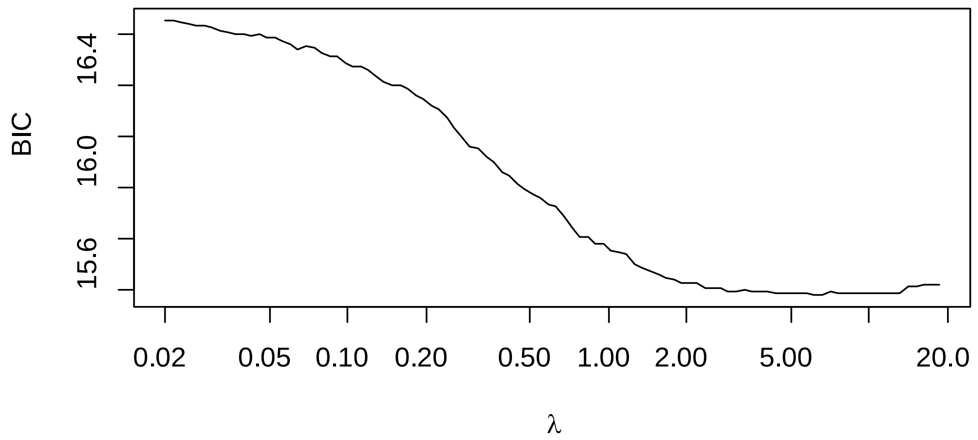


Figure 6.1: Adaptive group bridge with initial estimator from CRQR for 0.11 quantile
Note: CRQR refers to competing risks quantile regression.

6.2 Competing Risks Quantile Regression

For the competing risks quantile regression, we use variables that at least appear in Table 6.1 once and drop variables that are highly correlated. In the end we select 22 variables.

Of the 22 variables, we select 20 variables that are significant for at least some of the quantiles and show the estimated cumulative incidence curve in Figure 6.2 and estimated coefficients at a grid of quantiles that is equally spaced on $[\tau_L, \tau_U]$ with a step size of 0.01 in Figure 6.3. Apart from the detailed conditional distribution across quantiles from Figure 6.3, Table 6.2 shows some summary statistics for all the 22 variables, including trimmed mean effects and inference statistics of the constant test.

Figure 6.2 shows the estimated conditional cumulative incidence curve for a reference individual defined by setting all variables to sample averages.[*] According to the model, we cannot obtain estimates for the entire durations, but τ_U (0.58) is much lower than the share of observed transitions into retirement (86.12% in Table 4.1). One possible explanation is that for higher quantiles, that is longer durations, people enter into old age pension intensively at some time points. From the data, we can see a large mass for durations at several time points from week 260 to 310 (see Figure 4.1), so the share of transitions effective for analysis is actually lower than 86.12%. We can see a shaper increase in the middle quantiles, showing evidence that many people choose to defer early retirement pension for two years and then enter intensively at around 160 weeks from 2009.

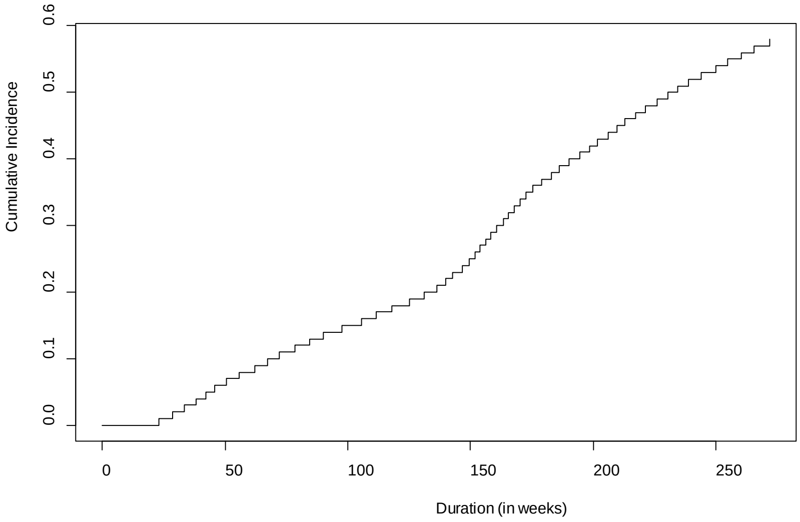
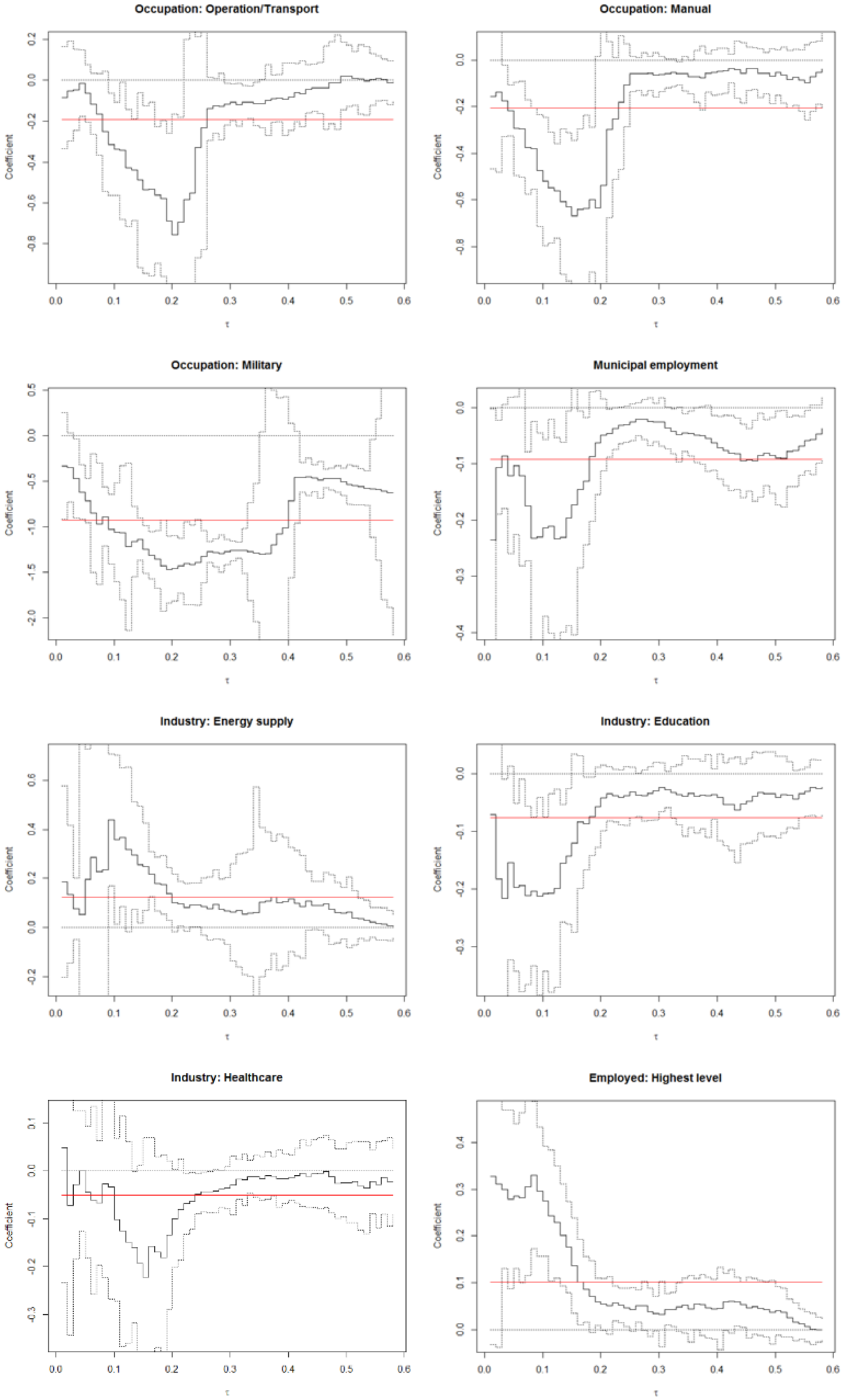


Figure 6.2: Estimated cumulative incidence curve for the reference individual

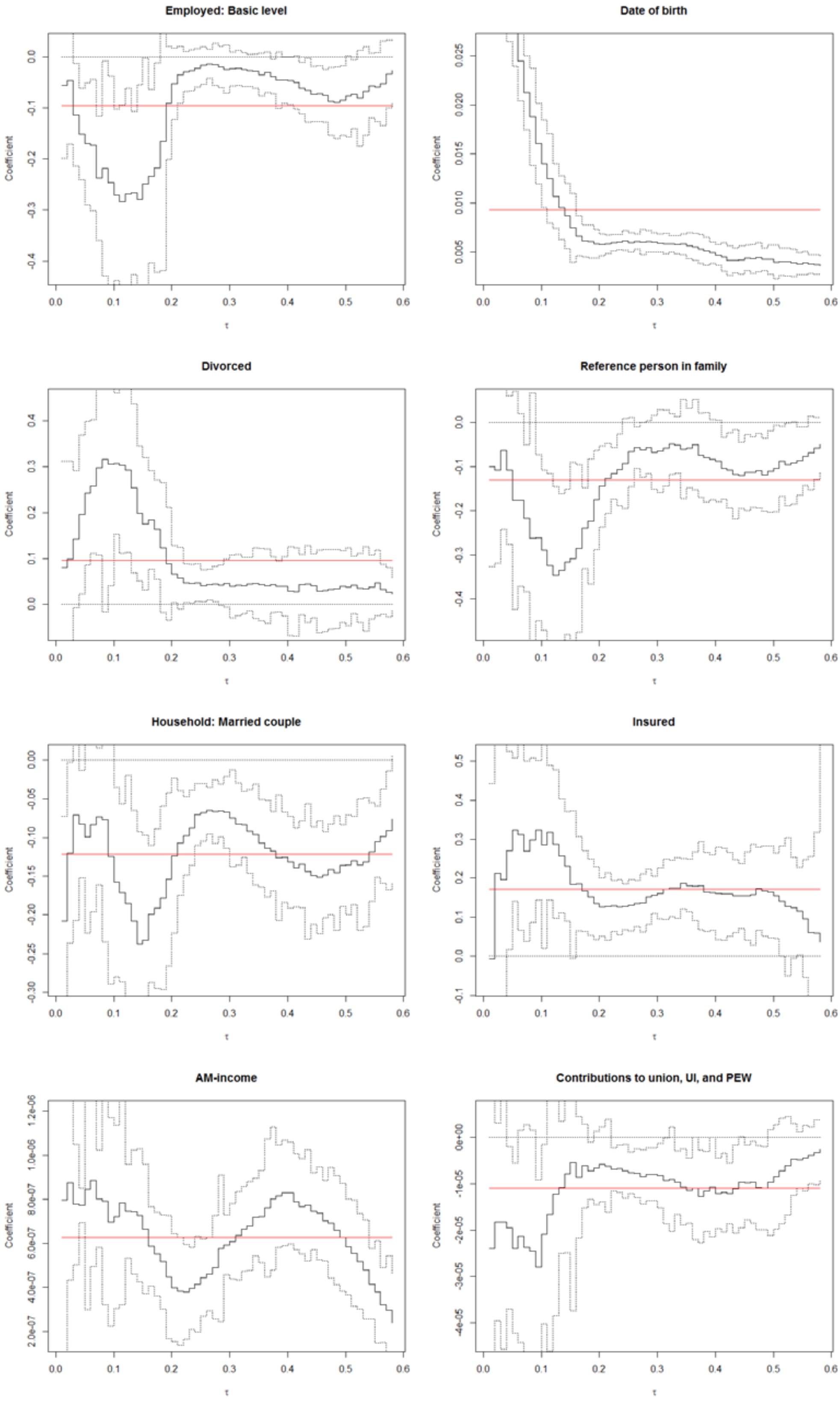
[*] This may not be the best approach given presence of dummy variables. But most patterns are similar across different specifications of the dummy variables for the reference individual, such as the sharper increase in the middle.

Figure 6.3 shows a detailed conditional distribution of estimated coefficients across quantiles. We can see that although the signs of all variables keep unchanged, the magnitude and significance level change across quantiles. The first two occupation variables are significant for lower quantiles, suggesting that these two occupations only increase the cumulative incidence, or decrease employment duration for people who retire the earliest, that is people who choose not to defer the early retirement pension for two years. The military occupation, on the contrary, have a negative and significant effect for almost all quantiles. Municipal employment and basic level employment have significant negative effects for both early retirement and old age pension, but are insignificant in the middle part where people who choose to defer early retirement pension. Being divorced and employed at the highest level have significant effects to decrease cumulative incidences for those who retire the earliest, but is insignificant afterwards. The variable ‘Date of birth’ has a decreasing positive effect over quantiles, suggesting that people who retire the earliest are most affected by the age eligibility of the PEW program. For people who retire later, it seems that more factors come into play and the role of age eligibility becomes smaller and smaller. Although it is unclear what the variable ‘insured’ means, it remains strong, positive and significant for almost all quantiles, so does the income variable. Although the magnitude of the coefficient of contributions to union, UI, and PEW does not change significantly across quantiles, we observe a drastic change for contributions to PEW. Similar to the variable ‘Date of birth’, contributions to PEW has a negative effect on observed transition probabilities that is decreasing in magnitude, which is reasonable given that for people who defer early retirement pension and enter old age pension, the contributions to early retirement scheme should matter less compared with those who retire the earliest. ATP contributions does not have a significant effect for most quantiles. Debt value has a strong positive and significant effect for most quantiles. Pension income is only significant for the lower and a small part of the middle quantile.



(Continued)

Figure 6.3: Selected estimates of the competing risks quantile regression



(Continued)

Figure 6.3: Selected estimates of the competing risks quantile regression (Continued)

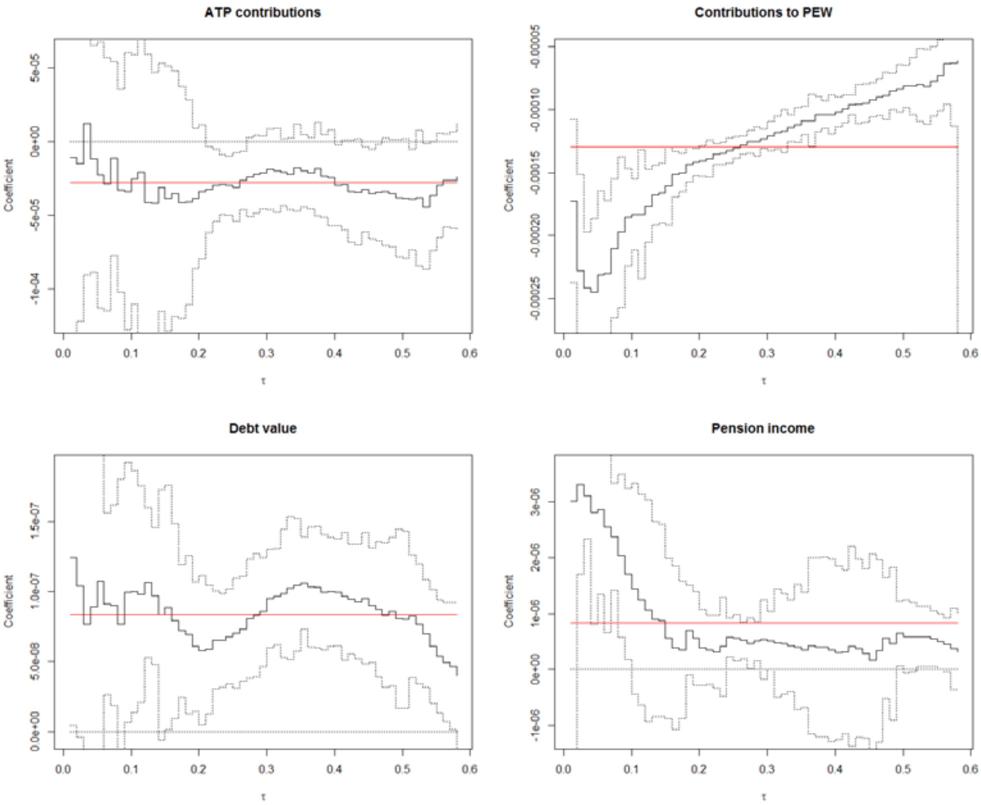


Figure 6.3: Selected estimates of the competing risks quantile regression (Continued)

Note: The solid staircase line refers to the estimated coefficients. The dotted staircase lines refer to the 95% asymptotic confidence intervals. The flat solid line refers to the trimmed mean effect. The flat dotted line refers to the value of zero.

Table 6.2 shows a summary of the trimmed mean effects and inference statistics of the constant test for the estimates. We can see that only three variables have insignificant trimmed mean effects and two of them are not included in Figure 6.3 due to being insignificant for almost all quantiles. 15 out of 19 variables are significant at 1% significance level and the rest are significant at 5% or 10% significance level, showing that most selected variables are of high quality and play important roles in transition from employment into retirement. The constant test shows that consistent with what we observe visually in Figure 6.3, 6 variables ‘occupation: operation/transport’, ‘occupation: manual’, ‘employed: highest level’, ‘date of birth’, ‘divorced’ and ‘contributions to PEW’ have highly significant test results, suggesting that the estimated coefficients vary significantly across quantiles. Half of the variables show constant effects on conditional quantiles, and the rest have test results that are significant at 5% or 10% significance level. These results highlight the relevance of competing risks quantile regression in providing detailed information on the heterogeneous effects of regressors on different cumulative incidence quantiles, in other words for people who enter different retirement programs, and thus could provide suggestions on late retirement policy targeted at different group of people.

Table 6.2: Trimmed mean effects and constant test results

Regi-sters	Within-register variables	Trimmed mean effect	SE of trimmed mean effect	P-value of trimmed mean effect	P-value of constant test
AKM	Occupation: Operation/Transport	-0.191	0.080	0.017	0.003
	Occupation: Manual	-0.206	0.060	0.001	0.000
	Occupation: Military	-0.932	0.172	0.000	0.080
	Municipal employment	-0.091	0.027	0.001	0.110
	Industry: Energy supply	0.124	0.073	0.089	0.190
	Industry: Education	-0.076	0.027	0.005	0.018
	Industry: Healthcare	-0.050	0.033	0.130	0.110
	Employed: Highest level	0.101	0.024	0.000	0.000
	Employed: Basic level	-0.095	0.029	0.001	0.022
BEF	Date of birth	0.009	0.001	0.000	0.000
	Divorced	0.096	0.031	0.002	0.002
	Reference person in family	-0.132	0.036	0.000	0.017
	Household: Married couple	-0.122	0.024	0.000	0.720
	Male	0.016	0.038	0.670	0.280
IDAP	Insured	0.172	0.038	0.000	0.270
IND	AM-income (million)	0.627	0.086	0.000	0.950
	Other capital income (million)	0.208	0.218	0.340	0.410
	ATP contributions (thousand)	-0.028	0.015	0.062	0.950
	Contributions to PEW (thousand)	-0.130	0.007	0.000	0.000
	Contributions to union, UI, and PEW (thousand)	-0.011	0.004	0.003	0.330
	Debt value (million)	0.084	0.018	0.000	0.980
	Pension income (million)	0.840	0.362	0.020	0.036

Chapter 7
Conclusions and Discussion

We adopt a flexible dependent competing risks setting to analyze the observed transitions from employment to retirement. We present how (adaptive) group bridge is helpful for bi-level variable selection in the competing risks quantile regression. To our knowledge, this is the first application of these methods to high-dimensional administrative data on the retirement problem. The data we use are register data and DREAM dataset provided by Statistics Denmark. Using (adaptive) group bridge, we select 4 registers and 28 within-register variables out of 16 registers and 397 within-register variables. The selected variables contain demographic, socioeconomic, financial and labor market information, have reasonable interpretation and also remain significant in the unpenalized competing risks quantile regression. From the competing risks quantile regression model, we find that the magnitude and significance level of estimated coefficients of most variables change significantly across quantiles, suggesting heterogeneous effects on transitions from employment into retirement for different durations and thus different retirement programs. Consistent with these results, in the (adaptive) group bridge, few variables are always selected over quantiles. These results suggest that the (adaptive) group bridge could reduce the dimension considerably and fit the competing risks quantile regression well, so it could be a promising method in the statistical toolbox for this type of problem.

Although we do not analyze the other exit routes out of employment, the results from one risk alone could provide evidence for the need of competing risks quantile regression. Thanks to the quantile regression technique, we do not face the problem Christensen and Kallestrup-Lamb (2012) find that lumping all retirement programs together could result in cancellation of opposite effects from certain variables on different retirement program. The reason is that in our sample, most people enter into either early retirement pension or old age pension, which are clearly separated in terms of durations, and the quantile regression technique can help us distinguish the heterogeneous effect of variables on different cumulative incidence quantiles and thus avoid cancellation.

However, we do face some other problems. The potential multicollinearity, measurement error and selection bias problems resulted from the data preparation process are already discussed in chapter 4, and the potential instability problem resulted from the BIC-type criterion is discussed in chapter 6.1. Another problem associated with multicollinearity is that we cannot fully understand the role of a register if the within-register variables that are important but highly correlated with variables from other registers are dropped. (Adaptive) group bridge allows the same variable to appear in different groups, but in our case, we have different but highly correlated variables, so additional work is needed to handle this problem.

Other questions and model extensions we can think of include: As Kallestrup-Lamb et al. (2016) point out, we should treat oracle property with caution. To what degree could we trust the selection results and how do those selected variables compare with conventional variables or variables from economic theory? As for model extensions, could we add an economic model at the bottom of competing risks quantile regression and will the results be different if we do so? Is it possible to allow for time-varying regressors and panel data structure? Is it possible to model the data generating process rather than the observed transition while using a flexible risks dependence specification? How to model unretirement in competing risks quantile regression given few observations?

Besides those open-ended questions, there are some practical improvements that are not too difficult to achieve: (1) We could add interaction terms to capture the heterogeneous effect of regressors among different wage groups, different gender, etc.; (2) We could add lags of regressors to capture dynamic effects; (3) We could add characteristics of spouses to capture the joint behavior of couples; (4) We could identify people who use unemployment as a pathway to retirement and include them in the retirement exit route; (5) We could identify people who choose to be semi-retired, that is to still work after entering into a retirement program and study separately transitions from employment to semi-retirement.

References

[1] Ahn, K. W., Banerjee, A., Sahr, N., & Kim, S. (2018). Group and within-group variable selection for competing risks data. Lifetime data analysis, 24(3), 407-424.

[2] Ahn, K., & Kim, S. (2018). Variable selection with group structure in competing risks quantile regression. Statistics in Medicine, 37(9), 1577-1586.

[3] Amilon, A., & Nielsen, T. H. (2010). How does the option to defer pension payments affect the labour supply of older workers in Denmark? In Working and Ageing: Emerging Theories and Empirical Perspectives (pp. 190-209). Thessaloniki: CEDEFOP - European Centre for the Development of Vocational Training.

[4] An, M., Christensen, B., & Gupta, N. (2004). Multivariate mixed proportional hazard modelling of the joint retirement of married couples. Journal of Applied Econometrics, 19(6), 687-704.

[5] Barslund, M. (2015). Extending working lives: The case of Denmark. CEPS working Document, (404).

[6] Bingley, & Lanot. (2007). Public pension programmes and the retirement of married couples in Denmark. Journal of Public Economics, 91(10), 1878-1901.

[7] Bingley, P., Gupta N.D. and Pedersen, P.J. (2012). Disability Programs, Health, and Retirement in Denmark since 1960. NBER Chapters,in: Social Security Programs and Retirement around the World: Historical Trends in Mortality and Health, Employment, and Disability Insurance Participation and Reforms, 217--249 National Bureau of Economic Research, Inc.

[8] Bingley, P., Gupta, N. D., & Pedersen, P. J. (2004). The Impact of Incentives on Retirement in Denmark. In J. Gruber, & D. Wise (Eds.), Social Security Programs and Retirement Around the World: Microestimation (pp. 153-234). Chicago: University of Chicago Press. A National Bureau of Economic Research conference report.

[9] Bingley, P., Gupta, N. D., Jørgensen, M., & Pedersen, P. J. (2016). Health, Disability Insurance, and Retirement in Denmark. In D. A. Wise (Ed.), Social Security Programs and Retirement around the World: Disability Insurance Programs and Retirement Chicago: University of Chicago Press.

[10] Breheny, P., & Huang, J. (2009). Penalized methods for bi-level variable selection. Statistics and its interface, 2(3), 369.

[11] Bühlmann, P., & Van De Geer, S. (2011). Statistics for high-dimensional data: methods, theory and applications. Springer Science & Business Media.

[12] Christensen, B., & Kallestrup-Lamb, M. (2012). The Impact of Health Changes on Labor Supply: Evidence from merged Data on Individual Objective MEdical Diagnosis Codes and Early Retirement Behavior. Health Economics, 21(Supp1), 56-100.

[13] Datta Gupta, N., & Larsen, M. (2007). Health shocks and retirement: The role of welfare state institutions. European Journal of Ageing, 4(3), 183-190.

[14] Datta Gupta, N., & Larsen, M. (2010). The impact of health on individual retirement plans: Self-reported versus diagnostic measures. Health Economics, 19(7), 792-813.

[15] Dlugosz, S. (2016). cmprskQR: Analysis of Competing Risks Using Quantile Regressions. R package version 0.9.1. https://CRAN.R-project.org/package=cmprskQR.

[16] Dlugosz, S., Lo, S., & Wilke, R. (2017). Competing risks quantile regression at work: In-depth exploration of the role of public child support for the duration of maternity leave. Journal of Applied Statistics, 44(1), 109-122.

[17] Duval, R. (2003). The Retirement Effects of Old-age Pension and Early Retirement Schemes in OECD Countries. OECD Economics DepartmentWorking Paper 370, OECD.

[18] Fan, J., & Li, R. (2001). Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. Journal of the American Statistical Association, 96(456), 1348-1360.

[19] Filges, T., Larsen, M. and Pedersen, P. J. (2012). Retirement: Does Individual Unemployment Matter? Evidence from Danish Panel Data 1980–2009. IZA Discussion Paper, No. 6538, Institute for the Study of Labor (IZA).

[20] Fine, J. P., & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. Journal of the American statistical association, 94(446), 496-509.

[21] Fitzenberger, B., & Wilke, R. (2010). Unemployment Durations in West Germany Before and After the Reform of the Unemployment Compensation System during the 1980s. German Economic Review, 11(3), 336-366.

[22] Fitzenberger, B., & Wilke, R. A. (2005). Using quantile regression for duration analysis. ZEW Discussion Paper, No. 05-65, Center for European Economic Research.

[23] Fitzenberger, B., & Wilke, R. A. (2015). Quantile Regression Methods. Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource, 1-18.

[24] Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of statistical software, 33(1), 1.

[25] Fu, W. J. (1998). Penalized regressions: the bridge versus the lasso. Journal of computational and graphical statistics, 7(3), 397-416.

[26] Gørtz, M. (2012). Early retirement in the day-care sector: The role of working conditions and health. European Journal of Ageing, 9(3), 187-198.

[27] Gruber, J., & Wise, D. (1998). Social Security and Retirement: An International Comparison. The American Economic Review, 88(2), 158-163.

[28] Hastie, T., Tibshirani, R., & Wainwright, M. (2015). Statistical learning with sparsity: the lasso and generalizations. Chapman and Hall/CRC.

[29] Hoerl, Arthur E. (1962). "Application of Ridge Analysis to Regression Problems". Chemical Engineering Progress, 58(3), 54–59.

[30] Huang, J., Breheny, P., & Ma, S. (2012). A Selective Review of Group Selection in High-Dimensional Models. Statistical Science, 27(4), 481-499.

[31] Huang, J., Liu, L., Liu, Y., & Zhao, X. (2014). Group selection in the Cox model with a diverging number of covariates. Statistica sinica, 1787-1810.

[32] Huang, J., Ma, S., Xie, H., & Zhang, C. (2009). A group bridge approach for variable selection. Biometrika, 96(2), 339-355.

[33] Kallestrup-Lamb, M., Kock, A., & Kristensen, J. (2016). Lassoing the Determinants of Retirement. Econometric Reviews, 35(8-10), 1-40.

[34] Kristensen, N. (2012). Training and retirement. IZA Discussion Paper, No. 6301, Institute for the Study of Labor (IZA).

[35] Kyyrä, T., & Wilke, R. (2007). Reduction in the Long-term Unemployment of the Elderly : A Success Story from Finland. Journal of the European Economic Association, 5(1), 154-182.

[36] Larsen, M., & J Pedersen, P. (2005). Pathways to Early Retirement in Denmark, 1984-2000. IZA Discussion Paper, No. 1575, Institute for the Study of Labor (IZA).

[37] Larsen, M., & Pedersen, P. (2013). To work, to retire – or both? Labor market activity after 60. IZA Journal of European Labor Studies, 2(1), 1-20.

[38] Lee, J. D., Sun, D. L., Sun, Y., & Taylor, J. E. (2016). Exact post-selection inference, with application to the lasso. The Annals of Statistics, 44(3), 907-927.

[39] Lindeboom, M. (1998). Microeconometric Analysis of the Retirement Decision: The Netherlands. OECD Economics Department Working Papers 207, OECD Publishing.

[40] Lockhart, R., Taylor, J., Tibshirani, R. J., & Tibshirani, R. (2014). A significance test for the lasso. Annals of statistics, 42(2), 413.

[41] Meinshausen, N., Meier, L., & Bühlmann, P. (2009). P-values for high-dimensional regression. Journal of the American Statistical Association, 104(488), 1671-1681.

[42] Miniaci, R. and Stancanelli, E. (1998). Microeconometric Analysis of the Retirement Decision: United Kingdom. OECD Economics Department Working Papers 206, OECD Publishing.

[43] Møller Danø, Ejrnæs, & Husted. (2005). Do single women value early retirement more than single men? Labour Economics, 12(1), 47-71.

[44] OECD (2012a). Thematic Follow-up Review of Policies to Improve Labour Market Prospects for Older Workers: Denmark. OECD Publishing.

[45] Peng, L., & Fine, J. (2009). Competing Risks Quantile Regression. Journal of the American Statistical Association, 104(488), 1440-1453.

[46] Peterson, A. V. (1976). Bounds for a joint distribution function with fixed sub-distribution functions: Application to competing risks. Proceedings of the National Academy of Sciences, 73(1), 11-13.

[47] Simon, N., Friedman, J., Hastie, T., & Tibshirani, R. (2013). A sparse-group lasso. Journal of Computational and Graphical Statistics, 22(2), 231-245.

[48] Taylor, J., & Tibshirani, R. J. (2015). Statistical learning and selective inference. Proceedings of the National Academy of Sciences, 112(25), 7629-7634.

[49] Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288.

[50] Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., & Knight, K. (2005). Sparsity and smoothness via the fused lasso. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(1), 91-108.

[51] Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 68(1), 49-67.

[52] Zhang, C. H. (2010). Nearly unbiased variable selection under minimax concave penalty. The Annals of statistics, 38(2), 894-942.

[53] Zhao, P., & Yu, B. (2006). On model selection consistency of Lasso. Journal of Machine learning research, 7(12), 2541-2563.

[54] Zhao, P., Rocha, G., & Yu, B. (2009). The composite absolute penalties family for grouped and hierarchical variable selection. The Annals of Statistics, 37(6A), 3468-3497.

[55] Zou, H. (2006). The Adaptive Lasso and Its Oracle Properties. Journal of the American Statistical Association, 101(476), 1418-1429.

[56] Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(2), 301-320.

Appendix

A.1 Tables

Table A.1: Descriptive statistics

	Sample		Right-censored		Retirement		Others	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Duration	200.3	101.0	418.0	0	204.3	94.52	126.6	81.02
AKM_atpsum2	1,176	121.9	1,184	122.0	1,176	123.5	1,171	108.8
AKM_discoalle_13_2	0.108	0.311	0.168	0.374	0.099	0.299	0.165	0.372
AKM_discoalle_13_5	0.111	0.314	0.047	0.212	0.115	0.319	0.095	0.294
AKM_discoalle_13_6	0.060	0.238	0.021	0.144	0.063	0.243	0.046	0.209
AKM_discoalle_13_8	0.042	0.200	0.026	0.16	0.043	0.203	0.037	0.189
AKM_discoalle_13_9	0.032	0.177	0.005	0.072	0.032	0.175	0.041	0.199
AKM_discoalle_13_10	0.062	0.241	0.021	0.144	0.065	0.246	0.049	0.215
AKM_discoalle_1	0.011	0.105	0.000	0.000	0.009	0.093	0.032	0.176
AKM_discoalle_3	0.282	0.450	0.529	0.500	0.278	0.448	0.266	0.442
AKM_discotype1	0.186	0.389	0.314	0.465	0.185	0.388	0.172	0.377
AKM_discotype2	0.361	0.480	0.257	0.438	0.374	0.484	0.288	0.453
AKM_discotype3	0.007	0.085	0.005	0.072	0.007	0.084	0.010	0.097
AKM_discotype4	0.441	0.497	0.424	0.496	0.430	0.495	0.526	0.500
AKM_funk_timeant	1,967	202.7	2,007	241.7	1,967	204.6	1,961	177.9
AKM_nace_5	0.008	0.090	0.005	0.072	0.009	0.092	0.005	0.073
AKM_nace_6	0.021	0.143	0.047	0.212	0.021	0.142	0.018	0.133
AKM_nace_7	0.085	0.279	0.042	0.201	0.083	0.275	0.111	0.315
AKM_nace_8	0.047	0.212	0.011	0.102	0.047	0.211	0.057	0.232
AKM_nace_12	0.009	0.097	0.000	0.000	0.010	0.099	0.008	0.092
AKM_nace_13	0.049	0.215	0.173	0.379	0.043	0.204	0.064	0.244
AKM_nace_13_4	0.012	0.108	0.016	0.125	0.012	0.107	0.013	0.112
AKM_nace_13_11	0.046	0.210	0.026	0.160	0.047	0.212	0.045	0.206
AKM_nace_13_14	0.028	0.166	0.021	0.144	0.028	0.165	0.031	0.173
AKM_nace_13_15	0.150	0.357	0.131	0.338	0.147	0.354	0.173	0.378
AKM_nace_16	0.190	0.392	0.230	0.422	0.193	0.395	0.155	0.362
AKM_nace_17	0.166	0.372	0.157	0.365	0.175	0.380	0.100	0.300
AKM_nace_18	0.020	0.139	0.016	0.125	0.020	0.142	0.014	0.117
AKM_nacea_3	0.099	0.299	0.068	0.253	0.096	0.295	0.128	0.334
AKM_nacea_10	0.044	0.205	0.031	0.175	0.042	0.200	0.064	0.244
AKM_nacea_19	0.020	0.139	0.016	0.125	0.021	0.144	0.010	0.097
AKM_nacei_1	0.942	0.235	0.874	0.332	0.946	0.225	0.918	0.274
AKM_nacei_2	0.015	0.122	0.031	0.175	0.014	0.117	0.020	0.141
AKM_nacei_14	0.011	0.103	0.011	0.102	0.010	0.098	0.019	0.137
AKM_socio13_4	0.283	0.450	0.529	0.500	0.278	0.448	0.269	0.444

	Sample		Right-censored		Retirement		Others	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
AKM_socio13_5	0.287	0.452	0.173	0.379	0.293	0.455	0.269	0.444
AKM_socio13_6	0.259	0.438	0.105	0.307	0.265	0.441	0.245	0.430
BEF_birthday	25.99	14.75	24.25	14.64	25.91	14.73	26.91	14.86
BEF_citizenship	0.989	0.104	0.995	0.072	0.989	0.103	0.986	0.117
BEF_civilstatus_1	0.763	0.425	0.743	0.438	0.763	0.425	0.770	0.421
BEF_civilstatus_2	0.033	0.178	0.031	0.175	0.034	0.181	0.024	0.154
BEF_civilstatus_3	0.128	0.334	0.141	0.349	0.128	0.334	0.128	0.334
BEF_familystatus	0.047	0.212	0.047	0.212	0.049	0.215	0.038	0.192
BEF_familytype_1	0.743	0.437	0.723	0.449	0.743	0.437	0.744	0.437
BEF_familytype_3	0.071	0.257	0.136	0.344	0.066	0.248	0.099	0.298
BEF_hustype_1	0.500	0.500	0.372	0.485	0.518	0.500	0.396	0.489
BEF_hustype_2	0.100	0.300	0.047	0.212	0.105	0.307	0.074	0.262
BEF_hustype_3	0.696	0.460	0.675	0.469	0.697	0.460	0.691	0.462
BEF_hustype_4	0.062	0.242	0.068	0.253	0.063	0.243	0.055	0.228
BEF_male	0.582	0.493	0.791	0.408	0.558	0.497	0.717	0.451
BEF_origin	0.972	0.164	0.974	0.160	0.973	0.162	0.967	0.178
BEF_region_1	0.101	0.301	0.168	0.374	0.096	0.295	0.122	0.327
BEF_region_2	0.100	0.299	0.089	0.285	0.099	0.298	0.109	0.312
BEF_region_3	0.105	0.307	0.188	0.392	0.101	0.302	0.118	0.322
BEF_region_4	0.008	0.088	0.011	0.102	0.008	0.089	0.006	0.080
BEF_region_5	0.057	0.233	0.058	0.234	0.056	0.230	0.068	0.252
BEF_region_6	0.105	0.307	0.094	0.293	0.105	0.306	0.111	0.315
BEF_region_7	0.081	0.272	0.068	0.253	0.082	0.274	0.073	0.260
BEF_region_8	0.124	0.330	0.052	0.223	0.128	0.335	0.109	0.312
BEF_region_9	0.146	0.353	0.136	0.344	0.148	0.355	0.137	0.344
BEF_region_10	0.071	0.258	0.063	0.243	0.073	0.259	0.065	0.246
BFL_atp_beloeb	2,441	688.4	2,397	651.3	2,436	692.6	2,484	662.4
BFL_bredt_beloeb (1K)	424.1	168.5	551.3	231.7	414.7	149.2	468.9	251.3
BFL_indberettede_timer	2,037	2,251	2,119	1,204	2,030	1,945	2,066	3,922
BFL_loentimer	1,973	201.1	2,016	236.3	1,973	203.2	1,965	174.6
BFL_smalt_beloeb (1K)	420.2	166.5	543.3	199.4	411.0	148.4	464.4	249.7
BFL_atpcode_1	0.720	0.449	0.749	0.435	0.714	0.452	0.763	0.426
BFL_atpcode_2	0.028	0.164	0.026	0.160	0.029	0.167	0.019	0.137
BFL_atpcode_3	0.331	0.470	0.429	0.496	0.334	0.472	0.288	0.453
BFL_atpcode_4	0.065	0.246	0.042	0.201	0.065	0.246	0.070	0.255
BFL_branche07_3	0.103	0.303	0.073	0.261	0.099	0.299	0.135	0.341
BFL_branche07_4	0.015	0.120	0.021	0.144	0.014	0.117	0.019	0.137
BFL_branche07_5	0.009	0.095	0.005	0.072	0.010	0.099	0.005	0.073
BFL_branche07_6	0.024	0.153	0.052	0.223	0.023	0.149	0.027	0.161
BFL_branche07_7	0.091	0.288	0.042	0.201	0.088	0.284	0.122	0.327
BFL_branche07_8	0.050	0.217	0.021	0.144	0.049	0.216	0.059	0.236
BFL_branche07_10	0.047	0.212	0.052	0.223	0.045	0.207	0.066	0.248
BFL_branche07_11	0.056	0.230	0.052	0.223	0.056	0.230	0.057	0.232
BFL_branche07_12	0.016	0.126	0.000	0.000	0.017	0.128	0.015	0.121
BFL_branche07_13	0.053	0.224	0.188	0.392	0.048	0.214	0.064	0.244

	Sample		Right-censored		Retirement		Others	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
BFL_branche07_14	0.040	0.196	0.042	0.201	0.041	0.197	0.036	0.186
BFL_branche07_15	0.179	0.383	0.230	0.422	0.176	0.381	0.192	0.394
BFL_branche07_16	0.215	0.411	0.298	0.459	0.217	0.412	0.183	0.387
BFL_branche07_17	0.173	0.379	0.157	0.365	0.183	0.386	0.106	0.308
BFL_branche07_18	0.022	0.146	0.021	0.144	0.022	0.148	0.018	0.133
BFL_branche07_19	0.041	0.198	0.037	0.188	0.043	0.204	0.021	0.144
BFL_fictitious_4	0.023	0.151	0.026	0.160	0.025	0.156	0.011	0.102
BFL_function_1	0.111	0.314	0.188	0.392	0.106	0.308	0.128	0.334
BFL_function_2	0.122	0.327	0.283	0.452	0.122	0.327	0.094	0.292
BFL_function_4	0.077	0.266	0.120	0.326	0.078	0.268	0.060	0.238
BFL_function_6	0.280	0.449	0.162	0.370	0.292	0.455	0.218	0.413
BFL_function_8	0.021	0.142	0.026	0.160	0.021	0.145	0.014	0.117
BFL_function_9	0.007	0.083	0.005	0.072	0.007	0.086	0.003	0.056
BFL_function_12	0.051	0.220	0.047	0.212	0.051	0.219	0.053	0.224
BFL_function_13	0.437	0.496	0.440	0.498	0.425	0.494	0.524	0.500
BFL_impute	0.200	0.400	0.288	0.454	0.195	0.396	0.222	0.416
BFL_incometype_3	0.007	0.085	0.005	0.072	0.008	0.087	0.004	0.065
BFL_province_1	0.202	0.402	0.335	0.473	0.193	0.395	0.239	0.427
BFL_province_2	0.140	0.347	0.141	0.349	0.138	0.345	0.153	0.360
BFL_province_3	0.067	0.250	0.141	0.349	0.063	0.243	0.082	0.274
BFL_province_4	0.007	0.085	0.005	0.072	0.008	0.086	0.005	0.073
BFL_province_5	0.043	0.204	0.063	0.243	0.043	0.204	0.040	0.197
BFL_province_6	0.091	0.288	0.063	0.243	0.092	0.289	0.091	0.288
BFL_province_7	0.083	0.275	0.084	0.278	0.084	0.277	0.072	0.259
BFL_province_8	0.134	0.341	0.063	0.243	0.138	0.345	0.123	0.328
BFL_province_9	0.149	0.356	0.178	0.384	0.150	0.357	0.142	0.349
BFL_province_10	0.080	0.271	0.068	0.253	0.081	0.273	0.071	0.257
BFL_province_11	0.104	0.306	0.084	0.278	0.108	0.310	0.082	0.274
BFL_sector_3	0.034	0.180	0.021	0.144	0.033	0.178	0.042	0.202
BFL_sector_4	0.013	0.115	0.021	0.144	0.014	0.117	0.007	0.086
BFL_sector_6	0.367	0.482	0.393	0.490	0.353	0.478	0.464	0.499
BFL_sector_7	0.044	0.205	0.037	0.188	0.045	0.206	0.042	0.202
BFL_sector_16	0.124	0.329	0.199	0.400	0.120	0.325	0.136	0.343
BFL_sector_17	0.106	0.308	0.262	0.441	0.105	0.307	0.079	0.271
BFL_sector_18	0.077	0.267	0.120	0.326	0.078	0.268	0.060	0.238
BFL_sector_20	0.283	0.450	0.162	0.370	0.294	0.456	0.220	0.415
BFL_sector_21	0.010	0.098	0.011	0.102	0.010	0.098	0.010	0.097
BFL_sector_23	0.010	0.098	0.005	0.072	0.009	0.096	0.013	0.112
BFL_sector_24	0.047	0.211	0.058	0.234	0.048	0.214	0.034	0.181
FAM_antb_1	0.021	0.181	0.131	0.512	0.018	0.162	0.021	0.183
FAM_antb_2	0.015	0.123	0.037	0.188	0.014	0.118	0.020	0.141
FAM_antb_3	0.043	0.214	0.115	0.336	0.041	0.212	0.038	0.192
FAM_antb_4	0.092	0.307	0.199	0.438	0.087	0.299	0.106	0.328
FAM_antpersf	1.976	0.647	2.267	0.972	1.967	0.633	1.986	0.662
FIDF_bagatel_1	0.991	0.095	0.974	0.160	0.991	0.095	0.994	0.080

	Sample		Right-censored		Retirement		Others	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
FIDF_bagatel_3	0.009	0.095	0.026	0.160	0.009	0.094	0.006	0.080
FIDF_branche07_3	0.097	0.295	0.068	0.253	0.094	0.292	0.123	0.328
FIDF_branche07_4	0.012	0.107	0.011	0.102	0.011	0.105	0.015	0.121
FIDF_branche07_6	0.020	0.141	0.052	0.223	0.020	0.139	0.017	0.129
FIDF_branche07_7	0.085	0.278	0.047	0.212	0.082	0.274	0.113	0.317
FIDF_branche07_8	0.046	0.210	0.011	0.102	0.046	0.209	0.056	0.230
FIDF_branche07_10	0.046	0.210	0.031	0.175	0.044	0.205	0.066	0.248
FIDF_branche07_11	0.047	0.211	0.026	0.160	0.047	0.213	0.045	0.206
FIDF_branche07_12	0.009	0.095	0.000	0.000	0.010	0.098	0.007	0.086
FIDF_branche07_13	0.044	0.204	0.157	0.365	0.039	0.193	0.058	0.234
FIDF_branche07_14	0.010	0.097	0.005	0.072	0.009	0.093	0.016	0.125
FIDF_branche07_15	0.442	0.497	0.346	0.477	0.453	0.498	0.383	0.486
FIDF_branche07_16	0.096	0.294	0.209	0.408	0.097	0.295	0.067	0.250
FIDF_branche07_17	0.015	0.121	0.011	0.102	0.016	0.123	0.011	0.102
FIDF_branche07_19	0.019	0.136	0.016	0.125	0.020	0.141	0.010	0.097
FIDF_funk_1	0.084	0.277	0.100	0.300	0.081	0.272	0.105	0.307
FIDF_funk_3	0.095	0.294	0.204	0.404	0.097	0.296	0.064	0.244
FIDF_funk_5	0.075	0.263	0.120	0.326	0.076	0.264	0.057	0.232
FIDF_funk_7	0.278	0.448	0.147	0.355	0.289	0.453	0.218	0.413
FIDF_funk_9	0.009	0.097	0.011	0.102	0.010	0.098	0.007	0.086
FIDF_funk_10	0.010	0.097	0.005	0.072	0.010	0.101	0.005	0.073
FIDF_funk_12	0.048	0.213	0.026	0.160	0.048	0.213	0.053	0.224
FIDF_funk_13	0.396	0.489	0.387	0.488	0.384	0.487	0.487	0.500
FIDF_gf_aarsv_2	6367	9378	6525	11182	6379	9288	6242	9652
FIDF_gf_ansatte_2 (1K)	7.819	11.67	7.886	13.56	7.848	11.57	7.592	12.03
FIDF_province_1	0.274	0.446	0.288	0.454	0.268	0.443	0.314	0.464
FIDF_province_3	0.077	0.266	0.052	0.223	0.080	0.271	0.057	0.232
FIDF_province_4	0.124	0.329	0.131	0.338	0.123	0.328	0.132	0.339
FIDF_province_5	0.075	0.264	0.178	0.384	0.072	0.259	0.079	0.271
FIDF_province_8	0.070	0.256	0.042	0.201	0.071	0.257	0.070	0.255
FIDF_province_9	0.056	0.230	0.047	0.212	0.056	0.231	0.054	0.226
FIDF_province_11	0.109	0.312	0.094	0.293	0.108	0.310	0.119	0.324
FIDF_region_2	0.176	0.381	0.141	0.349	0.178	0.383	0.168	0.374
FIDF_region_3	0.169	0.375	0.126	0.332	0.173	0.378	0.153	0.360
FIDF_region_4	0.478	0.500	0.602	0.491	0.468	0.499	0.529	0.499
FIDF_virkfkod_2	0.008	0.090	0.005	0.072	0.008	0.091	0.006	0.080
FIDF_virkfkod_4	0.349	0.477	0.335	0.473	0.337	0.473	0.442	0.497
FIDF_virkfkod_6	0.022	0.146	0.016	0.125	0.021	0.142	0.032	0.176
FIDF_virkfkod_7	0.009	0.097	0.011	0.102	0.010	0.098	0.007	0.086
FIDF_virkfkod_9	0.042	0.201	0.031	0.175	0.045	0.206	0.027	0.161
FIDF_virkfkod_11	0.010	0.098	0.005	0.072	0.010	0.098	0.011	0.102
FIDF_virkfkod_16	0.098	0.297	0.115	0.320	0.095	0.293	0.112	0.316
FIDF_virkfkod_17	0.074	0.262	0.120	0.326	0.075	0.264	0.057	0.232
FIDF_virkfkod_18	0.277	0.448	0.147	0.355	0.289	0.453	0.218	0.413
FIDF_virkfkod_20	0.093	0.290	0.204	0.404	0.093	0.291	0.065	0.246

	Sample		Right-censored		Retirement		Others	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
FIDF_year	21.55	23.92	22.66	28.13	21.27	23.49	23.39	25.99
FIRM_bd_1	0.634	0.482	0.639	0.482	0.648	0.478	0.525	0.500
FIRM_mk_2	0.327	0.469	0.335	0.473	0.312	0.463	0.440	0.497
IDAS_aarsvrk	398.9	898.1	591.2	1,296	386.5	868.2	452.9	1,007
IDAS_antaar	614.7	1,401	961.7	2,101	597.4	1,356	673.4	1,536
IDAS_antnov	464.4	1,065	703.9	1,558	451.1	1,032	515.6	1,175
IDAS_antnovbi	23.65	65.56	38.53	97.6	23.09	64.17	24.82	67.48
IDAS_branche_2	0.119	0.324	0.089	0.285	0.116	0.321	0.148	0.356
IDAS_branche_3	0.022	0.145	0.052	0.223	0.021	0.144	0.017	0.129
IDAS_branche_4	0.137	0.344	0.058	0.234	0.133	0.34	0.178	0.383
IDAS_branche_5	0.043	0.202	0.031	0.175	0.041	0.197	0.061	0.240
IDAS_branche_6	0.047	0.212	0.026	0.160	0.048	0.213	0.046	0.209
IDAS_branche_8	0.078	0.269	0.194	0.396	0.073	0.261	0.093	0.291
IDAS_branche_9	0.505	0.500	0.518	0.501	0.515	0.500	0.426	0.495
IDAS_branche_10	0.038	0.191	0.031	0.175	0.040	0.196	0.022	0.148
IDAS_filial_1	0.180	0.384	0.199	0.400	0.173	0.378	0.224	0.417
IDAS_filial_2	0.063	0.243	0.073	0.261	0.064	0.244	0.056	0.230
IDAS_filial_3	0.045	0.207	0.042	0.201	0.046	0.209	0.039	0.194
IDAS_filial_4	0.024	0.154	0.031	0.175	0.024	0.152	0.029	0.167
IDAS_filial_5	0.021	0.143	0.037	0.188	0.020	0.139	0.028	0.164
IDAS_filial_6	0.018	0.132	0.026	0.160	0.018	0.132	0.016	0.125
IDAS_filial_7	0.008	0.089	0.005	0.072	0.009	0.092	0.004	0.065
IDAS_filial_9	0.631	0.482	0.565	0.497	0.639	0.480	0.591	0.492
IDAS_idtilb_1	0.735	0.441	0.660	0.475	0.738	0.440	0.733	0.443
IDAS_idtilb_2	0.231	0.422	0.314	0.465	0.227	0.419	0.243	0.429
IDAS_idtilb_3	0.009	0.092	0.000	0.000	0.009	0.094	0.007	0.086
IDAS_idtilb_4	0.020	0.141	0.011	0.102	0.021	0.145	0.014	0.117
IDAN_ansdage	357.7	40.64	359.5	36.73	357.8	40.29	356.7	43.91
IDAN_ansxtilb_1	0.099	0.299	0.089	0.285	0.098	0.297	0.111	0.315
IDAN_ansxtilb_3	0.030	0.171	0.037	0.188	0.027	0.163	0.051	0.220
IDAN_ansxtilb_4	0.033	0.178	0.026	0.160	0.034	0.182	0.022	0.148
IDAN_ansxtilb_6	0.826	0.379	0.827	0.379	0.829	0.377	0.803	0.398
IDAN_type1	0.143	0.351	0.272	0.446	0.140	0.347	0.146	0.353
IDAN_type2	0.066	0.248	0.141	0.349	0.066	0.248	0.052	0.222
IDAN_type3	0.031	0.174	0.094	0.293	0.030	0.171	0.028	0.164
IDAN_type5	0.013	0.112	0.058	0.234	0.012	0.109	0.007	0.086
IDAN_year	10.26	9.332	9.895	8.917	10.45	9.411	8.969	8.703
IDAP_atpar	27.60	3.645	26.93	4.594	27.71	3.483	26.94	4.444
IDAP_ejnov	0.192	0.558	0.435	0.885	0.186	0.548	0.189	0.532
IDAP_erhver (1K)	26.30	4.267	25.76	5.119	26.36	4.144	25.92	4.914
IDAP_erhver79	7.163	3.462	5.545	3.524	7.211	3.435	7.133	3.571
IDAP_exit	0.068	0.253	0.089	0.285	0.065	0.247	0.087	0.282
IDAP_insured	0.947	0.224	0.874	0.332	0.963	0.189	0.842	0.365
IDAP_labyear	27.67	4.250	27.45	4.032	27.75	4.126	27.15	5.090
IDAP_member	0.942	0.233	0.874	0.332	0.962	0.192	0.813	0.391

	Sample		Right-censored		Retirement		Others	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
IDAP_memyear	22.81	7.776	20.92	9.107	23.42	7.300	18.70	9.452
IDAP_nsup	0.017	0.168	0.073	0.316	0.017	0.169	0.008	0.103
IDAP_pjob_dage	357.6	41.02	359.5	36.73	357.7	40.51	356.3	45.42
IDAP_ploentimer	1,889	231.8	1,900	208.4	1,889	227.0	1,887	269.0
IDAP_satp79	2,695	1,256	2,099	1,367	2,712	1,242	2,685	1,308
IDAP_sjob_dage	18.14	75.20	30.82	89.66	18.46	76.16	13.14	63.61
IDAP_sloentimer	29.23	174.8	35.52	120.8	30.12	179.3	21.36	148.0
IND_aekvivadisp_13 (1K)	314.4	144.4	393.2	299.1	311.1	136.5	322.9	148.8
IND_aindk94 (1K)	420.0	166.5	542.5	196.2	410.9	148.4	463.5	250.2
IND_andoverforsel	186.4	1,761	551.3	2,848	190.5	1,786	81.49	1,178
IND_ankapper (1K)	44.20	44.73	70.16	56.53	42.76	43.61	49.67	48.13
IND_arbfors (1K)	13.02	3.353	12.32	4.120	13.35	3.036	10.68	4.339
IND_askpli (1K)	0.248	4.293	0.160	2.175	0.223	2.635	0.451	10.34
IND_atpsaml	2,441	688.7	2,395	652.6	2,435	692.6	2,491	664.8
IND_bankakt (1K)	159.1	279.1	220.4	367.5	157.8	276.6	156.5	275.6
IND_bankgaeld (1K)	152.5	338.4	197.1	335.6	143.8	317.3	208.4	462.8
IND_befordr (1K)	6.084	11.89	6.382	12.48	5.945	11.74	7.055	12.82
IND_beskst13_3	0.045	0.208	0.100	0.300	0.041	0.198	0.066	0.248
IND_beskst13_4	0.954	0.209	0.895	0.307	0.959	0.199	0.933	0.250
IND_corfryns (1K)	4.168	14.81	6.281	19.37	3.852	13.79	6.104	20.06
IND_dispon_13 (1K)	276.7	144.7	392.8	379.1	270.9	129.3	296.3	156.3
IND_dispon_ny (1K)	281.8	147.3	402.4	383.4	275.7	131.6	302.4	159.6
IND_ejendom (1K)	1,390	1,730	2,550	5,440	1,330	1,510	1,580	1,690
IND_fagfkdb	4,077	2,180	4,252	2,612	4,137	2,141	3,599	2,309
IND_fosfufrd (1K)	9.336	27.13	18.11	104.1	8.893	21.92	10.87	24.31
IND_fradrag (1K)	84.59	85.94	120.1	174.5	81.76	81.28	98.54	89.82
IND_indbeeft	4,285	1,639	3,626	2,163	4,489	1,380	2,904	2,413
IND_kapindkp (1K)	-24.24	41.86	-34.19	96.66	-22.81	38.24	-32.95	47.78
IND_kapitialt (1K)	50.43	50.14	80.48	65.60	48.78	48.70	56.61	54.58
IND_kapitpriv (1K)	50.62	49.54	79.21	63.77	49.04	48.21	56.64	53.74
IND_koejd (1K)	1,380	1,770	2,520	4,970	1,320	1,490	1,620	2,300
IND_korydial	151.3	1,635	542.3	2,847	153.6	1,656	55.15	1,024
IND_kursakt (1K)	33.23	208.7	59.88	179.0	30.69	208.4	46.80	215.6
IND_lejev_egen_bolig (1K)	36.02	36.12	56.71	45.73	34.97	35.28	39.66	38.53
IND_lignfrdp (1K)	32.02	13.87	32.26	14.40	32.15	13.69	30.98	14.99
IND_loenmio	150.1	2,610	251.0	2,843	138.5	2,536	215.8	3,065
IND_loenmv (1K)	422.4	170.0	548.7	226.2	413.1	151.8	466.2	249.9
IND_loenskpl (1K)	421.2	166.4	543.7	196.1	412.0	148.4	465.1	249.5
IND_netovskud (1K)	5.252	55.78	36.15	262.6	4.178	34.32	7.013	63.92
IND_netovskud_13 (1K)	-0.030	49.69	18.66	200.4	-0.401	38.88	-1.043	44.26
IND_oblakt (1K)	44.09	203.0	87.54	366.1	42.53	196.4	46.98	204.6
IND_oblgaeld (1K)	448.6	660.9	806.3	1,500	422.1	601.1	574.3	767.2
IND_overforsindk (1K)	2.383	19.29	5.138	51.71	2.259	17.53	2.757	19.77
IND_perindkp (1K)	377.2	162.4	493.6	261.5	368.7	143.7	417.4	236.1

	Sample		Right-censored		Retirement		Others	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
IND_peroevrigformue (1K)	8.336	159.2	67.52	404.9	5.672	145.4	16.23	167.7
IND_qaktivf_ny05 (1K)	1,630	1,880	2,920	5,490	1,570	1,670	1,840	1,850
IND_qpassivn (1K)	608.5	870.8	1,010	1,590	573.5	819.3	788.0	989.0
IND_qpensialt (1K)	2.144	19.15	4.571	51.68	2.016	17.41	2.612	19.44
IND_qpenspri (1K)	1.021	14.83	4.571	51.68	0.917	12.74	1.079	12.44
IND_qrentud2 (1K)	36.37	54.59	60.03	79.09	34.06	50.76	48.83	70.97
IND_qtjpens (1K)	1.070	11.95	0.000	0.000	1.046	11.65	1.470	14.99
IND_rentbank (1K)	6.664	15.03	9.641	21.00	6.575	14.69	6.720	16.00
IND_rentudio	197.3	3,710	10.97	128.2	168.1	2,646	453.3	8,185
IND_rentupri (1K)	21.96	28.54	34.40	39.62	20.91	27.11	27.30	34.45
IND_rntiovir	160.4	3,306	751.1	4,897	122.4	3,037	323.9	4,575
IND_rudgbank (1K)	9.972	19.00	14.41	25.95	9.329	17.66	13.88	25.38
IND_samskat_1	0.237	0.425	0.262	0.441	0.237	0.425	0.233	0.423
IND_samskat_2	0.748	0.434	0.733	0.444	0.748	0.434	0.750	0.433
IND_skattot_13 (1K)	134.4	118.0	229.1	239.8	128.7	104.0	157.3	160.9
IND_sluskat (1K)	136.6	119.1	231.6	240.3	131.0	105.0	159.6	162.5
IND_virkkod_1	0.931	0.254	0.848	0.360	0.937	0.243	0.900	0.300
IND_virkkod_2	0.040	0.197	0.120	0.326	0.036	0.185	0.060	0.238
IND_virkordind (1K)	-0.413	46.89	14.24	120.5	-0.687	45.18	-1.331	29.23
INPI_arbpen10 (1K)	22.98	27.58	33.89	33.54	23.06	26.66	20.25	32.09
INPI_arbpen11 (1K)	17.63	49.19	24.37	49.26	16.60	45.46	23.96	70.72
INPI_arbpen12 (1K)	12.91	55.80	22.25	83.88	11.55	51.40	21.18	75.81
INPI_arbpen14 (1K)	5.188	10.53	4.795	11.44	5.198	10.44	5.199	11.07
INPI_arbpen15	1,223	5,506	569.0	2,728	1,227	5,555	1,321	5,554
INPI_arbpen16	690.5	1,901	1,427	2,839	680.8	1,866	614.3	1,891
INPI_pripen11 (1K)	1.734	14.89	3.171	12.85	1.594	10.22	2.488	33.31
INPI_pripen12 (1K)	6.761	37.49	14.29	103.9	6.161	21.43	9.718	81.01
INPI_pripen13	150.7	581.1	69.69	348.8	156.5	590.7	123.6	543.8
INPI_pripen15 (1K)	3.667	10.04	3.495	10.59	3.693	10.04	3.507	9.932
INPI_qpripen (1K)	13.84	42.51	21.96	104.6	13.05	27.18	18.10	88.96
INPI_qpripenl (1K)	8.989	40.66	17.73	104.1	8.168	24.17	13.35	88.01
KRIN_age	1.089	6.943	1.665	8.604	0.927	6.378	2.185	9.894
KRIN_crime	0.025	0.155	0.037	0.188	0.021	0.144	0.048	0.213
KRIN_length	2.880	121.9	37.91	521	2.002	97.03	2.338	59.4
KRIN_nocrime	0.035	0.256	0.063	0.418	0.029	0.226	0.070	0.387
KRIN_ubstrfko_1	0.009	0.095	0.011	0.102	0.008	0.090	0.015	0.121
KRIN_ubstrfko_2	0.018	0.131	0.026	0.160	0.015	0.120	0.037	0.189
KRIN_year8090	0.009	0.097	0.011	0.102	0.009	0.093	0.014	0.117
KRIN_year9000	0.014	0.116	0.031	0.175	0.012	0.107	0.027	0.161
LON_bfer (1K)	57.51	31.58	69.06	24.03	56.30	28.97	64.18	46.66
LON_bfra (1K)	9.020	9.822	9.280	17.10	8.842	9.279	10.30	11.52
LON_braarb_3	0.098	0.298	0.079	0.270	0.095	0.293	0.126	0.332
LON_braarb_4	0.014	0.115	0.016	0.125	0.013	0.112	0.018	0.133
LON_braarb_5	0.018	0.132	0.005	0.072	0.019	0.135	0.016	0.125
LON_braarb_6	0.022	0.147	0.052	0.223	0.022	0.146	0.018	0.133

	Sample		Right-censored		Retirement		Others	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LON_braarb_7	0.089	0.284	0.042	0.201	0.086	0.281	0.115	0.320
LON_braarb_8	0.047	0.212	0.011	0.102	0.047	0.211	0.057	0.232
LON_braarb_10	0.044	0.205	0.031	0.175	0.042	0.200	0.064	0.244
LON_braarb_11	0.048	0.214	0.052	0.223	0.049	0.215	0.045	0.206
LON_braarb_12	0.009	0.095	0.000	0.000	0.010	0.097	0.007	0.086
LON_braarb_13	0.047	0.212	0.147	0.355	0.043	0.202	0.060	0.238
LON_braarb_14	0.016	0.127	0.016	0.125	0.016	0.127	0.017	0.129
LON_braarb_15	0.157	0.364	0.136	0.344	0.155	0.362	0.178	0.383
LON_braarb_16	0.192	0.394	0.225	0.419	0.196	0.397	0.158	0.365
LON_braarb_17	0.175	0.380	0.162	0.370	0.185	0.388	0.110	0.313
LON_braarb_18	0.019	0.138	0.016	0.125	0.020	0.142	0.013	0.112
LON_braarb_19	0.020	0.141	0.016	0.125	0.022	0.146	0.011	0.102
LON_ferie_sh	41.00	24.97	49.27	23.94	40.14	23.54	45.71	33.31
LON_funk_1	0.120	0.325	0.173	0.379	0.108	0.311	0.196	0.397
LON_funk_2	0.282	0.450	0.545	0.499	0.277	0.448	0.267	0.443
LON_funk_3	0.292	0.455	0.178	0.384	0.297	0.457	0.271	0.445
LON_funk_4	0.111	0.314	0.047	0.212	0.115	0.319	0.095	0.294
LON_funk_5	0.063	0.242	0.021	0.144	0.066	0.248	0.047	0.211
LON_funk_7	0.041	0.197	0.026	0.160	0.042	0.200	0.036	0.186
LON_funk_8	0.033	0.178	0.005	0.072	0.032	0.176	0.043	0.204
LON_funk_9	0.075	0.264	0.021	0.144	0.078	0.268	0.065	0.246
LON_gene	4.734	16.12	6.580	21.93	4.712	15.88	4.525	16.56
LON_gtil (1K)	6.596	20.68	9.023	30.88	6.554	19.95	6.425	23.30
LON_gw	345.1	182.9	429.0	209.0	337.6	173.9	384.7	228.2
LON_jubgrat (1K)	1.352	24.47	4.933	62.46	1.142	21.09	2.198	32.90
LON_nw	237.2	118.1	297.3	146.6	232.6	112.4	259.3	144.6
LON_pens (1K)	75.81	78.37	92.69	65.42	73.39	72.62	90.43	112.7
LON_pension	53.32	55.48	64.39	45.45	51.69	52.28	63.20	75.64
LON_persgode (1K)	1.821	9.490	3.843	16.17	1.586	8.653	3.163	12.89
LON_pgod	2.754	14.23	5.728	23.69	2.414	13.04	4.691	19.16
LON_sector_1	0.458	0.498	0.450	0.499	0.446	0.497	0.547	0.498
LON_sector_2	0.188	0.391	0.319	0.467	0.187	0.390	0.174	0.379
LON_sector_3	0.362	0.481	0.257	0.438	0.375	0.484	0.288	0.453
LON_smftj (1K)	334.3	127.8	417.1	144.2	327.7	117.2	366.6	179.2
LON_timbet (1K)	94.72	1,220	312.5	1,990	97.62	1,250	29.07	659.8
LON_timferie	183.0	28.82	180.3	28.40	183.5	27.89	180.1	34.93
LON_timfra	45.57	49.71	39.83	87.94	45.31	46.85	48.64	58.83
LON_timover	17.72	71.94	21.74	69.91	17.59	71.39	17.84	76.30
LON_timprae	1,630	194.4	1,623	243.6	1,632	187.6	1,611	229.4
LON_timsh	61.54	10.00	60.35	12.37	61.62	9.661	61.13	11.81
LON_timuge	42.60	15.14	41.61	14.00	42.66	15.17	42.36	15.16
LON_ureglm (1K)	11.25	25.84	17.12	30.05	10.69	24.86	14.23	31.21
SGDP_antdage	0.299	6.071	4.471	24.55	0.107	2.679	0.886	11.85
SGDP_arbghp	113.3	2020	1334	6951	47.71	746.7	355.3	4594
SGDP_fraviaar	0.551	7.532	5.508	27.59	0.326	3.429	1.230	15.65

	Sample		Right-censored		Retirement		Others	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
SGDP_sagsart_1	0.016	0.125	0.052	0.223	0.015	0.121	0.016	0.125
SGDP_startsag_1	0.016	0.124	0.052	0.223	0.015	0.120	0.016	0.125
SYIN_diag23_3	0.011	0.104	0.026	0.160	0.010	0.099	0.016	0.125
SYIN_diag23_4	0.014	0.119	0.021	0.144	0.014	0.118	0.016	0.125
SYIN_diag23_6	0.010	0.098	0.016	0.125	0.009	0.096	0.013	0.112
SYIN_diag23_8	0.036	0.187	0.052	0.223	0.034	0.181	0.050	0.218
SYIN_diag99_11	0.011	0.103	0.016	0.125	0.010	0.100	0.014	0.117
SYIN_diag99_12	0.013	0.115	0.021	0.144	0.013	0.111	0.018	0.133
SYIN_diag99_15	0.020	0.139	0.026	0.160	0.019	0.138	0.022	0.148
SYIN_hosregion_1	0.009	0.095	0.016	0.125	0.009	0.093	0.012	0.107
SYIN_hosregion_2	0.009	0.095	0.005	0.072	0.009	0.096	0.008	0.092
SYIN_hosregion_3	0.011	0.103	0.037	0.188	0.010	0.099	0.013	0.112
SYIN_hosregion_6	0.009	0.096	0.011	0.102	0.009	0.096	0.010	0.097
SYIN_hosregion_8	0.008	0.092	0.005	0.072	0.009	0.093	0.007	0.086
SYIN_hosregion_9	0.011	0.104	0.011	0.102	0.010	0.099	0.019	0.137
SYIN_iantdg	0.222	1.174	0.330	1.306	0.204	1.047	0.329	1.846
SYIN_idiag_1	0.081	0.274	0.110	0.314	0.079	0.270	0.093	0.291
SYIN_idiag_2	0.010	0.098	0.016	0.125	0.009	0.095	0.013	0.112
SYIN_indm_1	0.056	0.230	0.089	0.285	0.054	0.226	0.068	0.252
SYIN_indm_2	0.033	0.179	0.037	0.188	0.033	0.177	0.038	0.192
SYIN_kapitlnr_2	0.011	0.104	0.021	0.144	0.010	0.101	0.014	0.117
SYIN_year_1	0.043	0.203	0.042	0.201	0.042	0.200	0.053	0.224
SYIN_year_2	0.045	0.207	0.068	0.253	0.043	0.204	0.050	0.218
UDDA_eduarea_1	0.142	0.349	0.047	0.212	0.143	0.350	0.148	0.356
UDDA_eduarea_2	0.026	0.158	0.031	0.175	0.025	0.155	0.032	0.176
UDDA_eduarea_3	0.096	0.295	0.037	0.188	0.098	0.297	0.098	0.297
UDDA_eduarea_4	0.034	0.180	0.110	0.314	0.032	0.175	0.033	0.178
UDDA_eduarea_7	0.032	0.175	0.058	0.234	0.030	0.172	0.037	0.189
UDDA_eduarea_8	0.240	0.427	0.126	0.332	0.245	0.430	0.227	0.419
UDDA_eduarea_9	0.015	0.120	0.052	0.223	0.014	0.117	0.014	0.117
UDDA_eduarea_11	0.095	0.293	0.115	0.320	0.091	0.288	0.114	0.318
UDDA_eduarea_12	0.078	0.268	0.084	0.278	0.076	0.265	0.090	0.286
UDDA_eduarea_13	0.061	0.238	0.131	0.338	0.060	0.237	0.051	0.220
UDDA_eduarea_14	0.011	0.105	0.031	0.175	0.011	0.102	0.012	0.107
UDDA_eduarea_15	0.139	0.346	0.152	0.360	0.145	0.352	0.093	0.291
UDDA_eduarea_16	0.009	0.093	0.011	0.102	0.009	0.094	0.006	0.080
UDDA_edulevel_2	0.141	0.348	0.047	0.212	0.143	0.350	0.147	0.355
UDDA_edulevel_3	0.412	0.492	0.236	0.425	0.420	0.494	0.391	0.488
UDDA_edulevel_4	0.047	0.212	0.021	0.144	0.046	0.210	0.060	0.238
UDDA_edulevel_5	0.271	0.444	0.277	0.449	0.274	0.446	0.245	0.430
UDDA_edulevel_6	0.120	0.325	0.382	0.487	0.109	0.312	0.146	0.353
UDDA_inst_1	0.009	0.093	0.011	0.102	0.008	0.088	0.016	0.125
UDDA_inst_3	0.007	0.084	0.016	0.125	0.007	0.085	0.004	0.065
UDDA_inst_4	0.047	0.212	0.178	0.384	0.044	0.204	0.046	0.209
UDDA_inst_7	0.012	0.110	0.031	0.175	0.010	0.099	0.027	0.161

	Sample		Right-censored		Retirement		Others	
Variables	Mean	SD	Mean	SD	Mean	SD	Mean	SD
UDDA_inst_10	0.015	0.120	0.052	0.223	0.013	0.113	0.019	0.137
UDDA_inst_13	0.008	0.087	0.016	0.125	0.008	0.087	0.005	0.073
UDDA_inst_18	0.009	0.093	0.016	0.125	0.008	0.090	0.013	0.112
UDDA_inst_20	0.028	0.165	0.068	0.253	0.025	0.157	0.039	0.194
UDDA_inst_26	0.514	0.500	0.733	0.444	0.507	0.500	0.520	0.500

Table A.2: Definition and links of selected variables

Variables	Definition
Occupations	Danish occupation code category 8, 9, 0 and 2
https://www.dst.dk/da/Statistik/dokumentation/nomenklaturer/disco-08	
Municipal employment	Source of employee occupation code category 2
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/discotyp	
Industries	Industry code category D, P and Q
https://www.dst.dk/da/Statistik/dokumentation/nomenklaturer/dansk-branchekode-db07	
Socioeconomic classifications	Socioeconomic classification code category 132, 134
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/socio13	
Family status	A person's position in the family category 1: reference person
https://www.dst.dk/da/Statistik/dokumentation/Times/forebyggelsesregistret/plads	
Household type	Household composition type category 3: A married couple
https://www.dst.dk/da/Statistik/dokumentation/Times/moduldata-for-befolkning-og-valg/hustype	
Insured	Insurance code category D and H
https://www.dst.dk/da/Statistik/dokumentation/Times/moduldata-for-arbejdsmarked/forsikringskategori-kode	
AM-income (million)	Labor market contribution-related income
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/aindk94	
Capital income (million)	Total capital income excluding equity income
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/kapitialt	
Other capital income (million)	
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/ankapper	
ATP contributions (thousand)	
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/atpsaml	
Contributions to PEW (thousand)	
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/indbeeft	
Property value (million)	
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/koejd	
Debt value (million)	Market value of bond debt
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/oblgaeld	
Pension income (million)	
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/qpensialt	
Interest expense (million)	Interest expense relating to mortgage debt
https://www.dst.dk/da/Statistik/dokumentation/Times/personindkomst/rentupri	

A.2 R Code

```
library(cmprskQR);library(quantreg);library(survival);library(haven);library(dplyr);library(muhaz);library(MASS);library(Matrix)

### Initial Estimator from Competing Risks Quantile Regression

tempdata<-cbind(time=ftime,cause=fstatus,X)
ftime<-tempdata$time
fstatus<-tempdata$cause
X<-tempdata[,3:400]
sds<-apply(X,2,sd)
means<-apply(X,2,mean)
X<-t((t(X)-means)/sds)
X<-cbind(rep(1,dim(X)[1]),X)
cvt.length<-ncol(X)
num<-nrow(X)
cencode<-0
outcome<-1
FT1<-as.numeric(fstatus==outcome)
FT1.csurv.x<-FT1
n.cvt.1<-crossprod(-X,FT1.csurv.x)
X.FT1.csurv.x<-X*FT1.csurv.x

tol<-0.0001
gamma<-0.5
cj<-c(1,rep(36^(1-gamma),36),rep(24^(1-gamma),24),rep(58^(1-gamma),58),rep(5^(1-gamma),5),rep(47^(1-gamma),47),rep(2^(1-gamma),2),rep(24^(1-gamma),24),rep(10^(1-gamma),10),rep(15^(1-gamma),15),rep(56^(1-gamma),56),rep(12^(1-gamma),12),rep(8^(1-gamma),8),rep(48^(1-gamma),48),rep(5^(1-gamma),5),rep(21^(1-gamma),21),rep(27^(1-gamma),27))
curr.tau<-0.11
n.cvt.2<-apply(2*X*curr.tau,2,sum)
max.lmt<-max(c(abs(n.cvt.1),abs(n.cvt.2)))+10000
pseudo.resp<-c(ftime*FT1.csurv.x,max.lmt,max.lmt)
pseudo.cvt<-rbind(X.FT1.csurv.x,as.vector(n.cvt.1),as.vector(n.cvt.2))
fit<-rq.fit.fnb(pseudo.cvt,pseudo.resp,tau = 0.5)

### Group Bridge Estimator

bic<-rep(0,length(Lambda))
pnum<-rep(0,length(Lambda))
```

```
betabridge<-matrix(0,length(initial),length(Lambda))
BGB<-matrix(0,length(tau.seq),length(initial))

Lambda<-exp(seq(log(0.05),log(50),by=(log(50)-log(0.05))/100)[-101])
initial<-fit$coefficient
for(j in 1:length(Lambda)){
  betapost<-initial;betapre<-rep(100,length(initial));i1<-1
  while(sum(abs(betapost-betapre))>0.001){
    betapre<-betapost;weightbeta<-rep(0,length(initial));
    weightbeta[1]<-abs(betapost[1])^gamma
    weightbeta[2:37]<-sum(abs(betapost[2:37]))^gamma
    weightbeta[38:61]<-sum(abs(betapost[38:61]))^gamma
    weightbeta[62:119]<-sum(abs(betapost[62:119]))^gamma
    weightbeta[120:124]<-sum(abs(betapost[120:124]))^gamma
    weightbeta[125:171]<-sum(abs(betapost[125:171]))^gamma
    weightbeta[172:173]<-sum(abs(betapost[172:173]))^gamma
    weightbeta[174:197]<-sum(abs(betapost[174:197]))^gamma
    weightbeta[198:207]<-sum(abs(betapost[198:207]))^gamma
    weightbeta[208:222]<-sum(abs(betapost[208:222]))^gamma
    weightbeta[223:278]<-sum(abs(betapost[223:278]))^gamma
    weightbeta[279:290]<-sum(abs(betapost[279:290]))^gamma
    weightbeta[291:298]<-sum(abs(betapost[291:298]))^gamma
    weightbeta[299:346]<-sum(abs(betapost[299:346]))^gamma
    weightbeta[347:351]<-sum(abs(betapost[347:351]))^gamma
    weightbeta[352:372]<-sum(abs(betapost[352:372]))^gamma
    weightbeta[373:399]<-sum(abs(betapost[373:399]))^gamma
    weight1<-(((1-gamma)/gamma)^gamma*cj*weightbeta)^(1-1/gamma)*cj^(1/gamma)*Lambda[j]
    fit.lasso<-rq.fit.lasso(pseudo.cvt,pseudo.resp,tau=0.5,lambda=weight1)
    betapost<-fit.lasso$coefficient;i1<-i1+1
    if(i1>200)break
  }
  betapost[abs(betapost)<tol]<-0;betabridge[,j]<-betapost;
  betabridge[-1,j]<-betapost[-1]/sds
  pnum[j]<-sum(1*(betapost!=0))
  bic[j]<-checkrev(pseudo.resp,pseudo.cvt,betapost)
}
bic1<-2*bic/num+pnum*log(num)/num/2*log(length(initial))
choice<-(1:length(Lambda))[bic1==min(bic1)][1];
BGB<-betabridge[,choice]

### Adaptive Group Bridge with Group Bridge as Initial Estimator

betaSageGB<-matrix(0,length(initial),length(Lambda))
```

```
for(j in 1:length(Lambda)){
  initial<-betabridge[,j]+0.00001;adapweight<-abs(betabridge[,j])+0.00001;
  betapost<-initial;betapre<-rep(100,length(initial));i1<-1
  while(sum(abs(betapost-betapre))>0.001){
    betapre<-betapost;weightbeta<-rep(0,length(initial));
    weightbeta[1]<-abs(betapost[1]/adapweight[1])^gamma
    weightbeta[2:37]<-sum(abs(betapost[2:37]/adapweight[2:37]))^gamma
    weightbeta[38:61]<-sum(abs(betapost[38:61]/adapweight[38:61]))^gamma
    weightbeta[62:119]<-sum(abs(betapost[62:119]/adapweight[62:119]))^gamma
    weightbeta[120:124]<-sum(abs(betapost[120:124]/adapweight[120:124]))^gamma
    weightbeta[125:171]<-sum(abs(betapost[125:171]/adapweight[125:171]))^gamma
    weightbeta[172:173]<-sum(abs(betapost[172:173]/adapweight[172:173]))^gamma
    weightbeta[174:197]<-sum(abs(betapost[174:197]/adapweight[174:197]))^gamma
    weightbeta[198:207]<-sum(abs(betapost[198:207]/adapweight[198:207]))^gamma
    weightbeta[208:222]<-sum(abs(betapost[208:222]/adapweight[208:222]))^gamma
    weightbeta[223:278]<-sum(abs(betapost[223:278]/adapweight[223:278]))^gamma
    weightbeta[279:290]<-sum(abs(betapost[279:290]/adapweight[279:290]))^gamma
    weightbeta[291:298]<-sum(abs(betapost[291:298]/adapweight[291:298]))^gamma
    weightbeta[299:346]<-sum(abs(betapost[299:346]/adapweight[299:346]))^gamma
    weightbeta[347:351]<-sum(abs(betapost[347:351]/adapweight[347:351]))^gamma
    weightbeta[352:372]<-sum(abs(betapost[352:372]/adapweight[352:372]))^gamma
    weightbeta[373:399]<-sum(abs(betapost[373:399]/adapweight[373:399]))^gamma
    weight1<-(((1-gamma)/gamma)^gamma*cj*weightbeta)^(1-
1/gamma)*cj^(1/gamma)/abs(adapweight)*Lambda[j]
    fit.lasso<-rq.fit.lasso(pseudo.cvt,pseudo.resp,tau=0.5,lambda=weight1)
    betapost<-fit.lasso$coefficient;i1<-i1+1
    if(i1>200)break
  }
  betapost[abs(betapost)<tol]<-0;betaSageGB[,j]<-betapost;
  betaSageGB[-1,j]<-betapost[-1]/sds
  pnum[j]<-sum(1*(betapost!=0))
  bic[j]<-checkrev(pseudo.resp,pseudo.cvt,betapost)
}
bic1<-2*bic/num+pnum*log(num)/num/2*log(length(initial))
choice<-(1:length(Lambda))[bic1==min(bic1)][1];
BSageGB<-betaSageGB[,choice]

### Adaptive Group Bridge with Initial Estimator from Competing Risks Quantile Regression

Lambda<-exp(seq(log(0.02),log(20),by=(log(20)-log(0.02))/100)[-101])
betaSageQ<-matrix(0,length(initial),length(Lambda))
initial<-fit$coefficients;adapweight<-abs(fit$coefficients);
for(j in 1:length(Lambda)){
  betapost<-initial;betapre<-rep(100,length(initial));i1<-1
  while(sum(abs(betapost-betapre))>0.001){
    betapre<-betapost;weightbeta<-rep(0,length(initial));
    weightbeta[1]<-abs(betapost[1]/adapweight[1])^gamma
    weightbeta[2:37]<-sum(abs(betapost[2:37]/adapweight[2:37]))^gamma
    weightbeta[38:61]<-sum(abs(betapost[38:61]/adapweight[38:61]))^gamma
    weightbeta[62:119]<-sum(abs(betapost[62:119]/adapweight[62:119]))^gamma
    weightbeta[120:124]<-sum(abs(betapost[120:124]/adapweight[120:124]))^gamma
    weightbeta[125:171]<-sum(abs(betapost[125:171]/adapweight[125:171]))^gamma
    weightbeta[172:173]<-sum(abs(betapost[172:173]/adapweight[172:173]))^gamma
    weightbeta[174:197]<-sum(abs(betapost[174:197]/adapweight[174:197]))^gamma
    weightbeta[198:207]<-sum(abs(betapost[198:207]/adapweight[198:207]))^gamma
    weightbeta[208:222]<-sum(abs(betapost[208:222]/adapweight[208:222]))^gamma
    weightbeta[223:278]<-sum(abs(betapost[223:278]/adapweight[223:278]))^gamma
    weightbeta[279:290]<-sum(abs(betapost[279:290]/adapweight[279:290]))^gamma
    weightbeta[291:298]<-sum(abs(betapost[291:298]/adapweight[291:298]))^gamma
    weightbeta[299:346]<-sum(abs(betapost[299:346]/adapweight[299:346]))^gamma
    weightbeta[347:351]<-sum(abs(betapost[347:351]/adapweight[347:351]))^gamma
    weightbeta[352:372]<-sum(abs(betapost[352:372]/adapweight[352:372]))^gamma
    weightbeta[373:399]<-sum(abs(betapost[373:399]/adapweight[373:399]))^gamma
    weight1<-(((1-gamma)/gamma)^gamma*cj*weightbeta)^(1-
1/gamma)*cj^(1/gamma)/abs(adapweight)*Lambda[j]
    fit.lasso<-rq.fit.lasso(pseudo.cvt,pseudo.resp,tau=0.5,lambda=weight1)
    betapost<-fit.lasso$coefficient;i1<-i1+1
    if(i1>200)break
  }
  betapost[abs(betapost)<tol]<-0;betaSageQ[,j]<-betapost;
  betaSageQ[-1,j]<-betapost[-1]/sds
  pnum[j]<-sum(1*(betapost!=0))
  bic[j]<-checkrev(pseudo.resp,pseudo.cvt,betapost)
}
bic1<-2*bic/num+pnum*log(num)/num/2*log(length(initial))
choice<-(1:length(Lambda))[bic1==min(bic1)][1];
BSageQ<-betaSageQ[,choice]

checkrev<-function(yy,xx,betatemp){
  tempsum<-crossprod(t(xx),betatemp)
  sum1<-sum(abs(yy-tempsum))
  return(sum1)
}

### Revised Code for R-Package
```

```
crrQR.int.new <- function(ftime, fstatus, X, tau.L, tau.U, tau.step,   outcome = 1,
                            cencode = 0,   noconst=FALSE, variance=TRUE, offset=0, max.lmt=10^8,
                            rq.method="br", orig.num=nrow(X), ...){

  eddcmp <- function(M){
    ev <- eigen(M)
    return(list(evectors=ev$vectors, evalues=ev$values, im.evalues=NULL))
  }

  if(!noconst){
    X <- cbind(rep.int(1, dim(X)[1]), X)
    dimnames(X)[[2]][1] <- "const"
  }

  cvt.length <- ncol(X)
  tau.seq <- seq(tau.L, tau.U, tau.step)
  L.tau.seq <- length(tau.seq)
  num <- nrow(X)
  FT0 <- as.numeric(fstatus==cencode)
  cens <- (sum(FT0)>0)
  FT1 <- as.numeric(fstatus==outcome)
  smallest <- .Machine$double.eps^0.5
  FT1.csurv.x <- FT1
  pseudo.resp   <- c((ftime-offset)*FT1.csurv.x, max.lmt, max.lmt)
  n.cvt.1 <- crossprod(-X, FT1.csurv.x)
  est.beta.seq <- NULL

  if(variance){
    est.var.seq <- NULL
    inf.est.func <- list()
    ind.conv.var <- rep(1, cvt.length)
    est.var <- rep(0, cvt.length)
    curr.inf <- array(rep(0, num*(cvt.length)), c(num, cvt.length))
    compare.vec <- rank(-ftime, ties.method="max")
    compare.mat <- (ftime>=t(array(rep(ftime,num), c(num, num))))*1
  }

  L.bsq <- 1
  ind.conv <- 1
  X.FT1.csurv.x <- X*FT1.csurv.x

  goon <- TRUE
  while(goon){
    est.beta <- rep(0, cvt.length)
    curr.tau <- tau.seq[L.bsq]
    n.cvt.2 <- apply(2*X*curr.tau, 2, sum)
    pseudo.cvt <- rbind(X.FT1.csurv.x, as.vector(n.cvt.1), as.vector(n.cvt.2))
    est.beta.obj <- rq.fit.br(pseudo.cvt, pseudo.resp, ci=FALSE)
    est.beta <- est.beta.obj$coef
    ind.conv<-
as.numeric(abs(est.beta.obj$residuals[num+1L])>1)*as.numeric(abs(est.beta.obj$residuals[num+2])>1)

    if(ind.conv==1){
      est.beta.seq <- rbind(est.beta.seq, est.beta)

      if(variance){
        helper <- as.numeric((ftime) < X %*% as.vector(est.beta)) * FT1.csurv.x
        tmp.1 <- ((helper-curr.tau)*X)
        var.matrix.cmp.1 <- crossprod(tmp.1, tmp.1)/num
        ZI.mat <- X*helper
        tmp.2.1 <- t(t(ZI.mat)%*%compare.mat)
        tmp.2 <- FT0*(tmp.2.1/compare.vec)
        var.matrix.cmp.2 <- crossprod(tmp.2, tmp.2)/num
        var.matrix <- var.matrix.cmp.1-var.matrix.cmp.2
        var.dcmp <- eddcmp(var.matrix)
        sigma.sqrt <- (var.dcmp$evector%*%diag(sqrt(var.dcmp$evalues),
                                               length(var.dcmp$evalues),
                                               length(var.dcmp$evalues))%*%solve(var.dcmp$evector))
        est.beta.var <- NULL
        ind.conv.var <- NULL

        for(k in 1:cvt.length){
          pseudo.cvt.var  <-  rbind(X.FT1.csurv.x,  as.vector(n.cvt.1),  as.vector(n.cvt.2)+2*sigma.sqrt[,
k]*sqrt(num))
          est.beta.var.obj <- rq.fit.br(pseudo.cvt.var, pseudo.resp, ci=FALSE)
          ind.conv.var <- c(ind.conv.var, as.numeric(abs(est.beta.var.obj$residuals[num+1])>1)*
                            as.numeric(abs(est.beta.var.obj$residuals[num+2])>1))
          est.beta.var <- cbind(est.beta.var, est.beta.var.obj$coef-est.beta)
        }

        if(sum(ind.conv.var)==cvt.length){
          est.var <- diag(tcrossprod(est.beta.var, est.beta.var))
          est.inv.deriv.matrx <- est.beta.var%*%solve(sigma.sqrt)*sqrt(num)
          est.inv.deriv.matrx <- (est.inv.deriv.matrx+t(est.inv.deriv.matrx))/2
          curr.inf <- tcrossprod(tmp.1-tmp.2, est.inv.deriv.matrx)
        } else{
```

```
        goon <- FALSE
      }
      inf.est.func[[L.bsq]] <- curr.inf
      est.var.seq <- rbind(est.var.seq, est.var)
    }
    L.bsq <- L.bsq+1
  } else{
    goon <- FALSE
  }

  goon <- (L.bsq<=L.tau.seq & goon)

  cat('.')
}
cat('done.\n')

if(L.bsq<L.tau.seq){
  cat('\n')
  cat('stopping at tau=')
  cat(tau.seq[L.bsq])
  cat('\n')
  L.tau.seq <- L.bsq-1
  if(L.bsq<=1) return(list(tau.seq=NULL, L.bsq=0))
  tau.seq <- tau.seq[1:L.tau.seq]
  tau.U <- tau.seq[L.tau.seq]
}

dimnames(est.beta.seq)[[1]] <- tau.seq
if(variance) dimnames(est.var.seq)[[1]] <- tau.seq

ret <- list(call=match.call(),
          beta.seq=est.beta.seq,
          tau.seq=tau.seq,
          cvt.length=cvt.length,
          n=num,
          n.missing=orig.num-num)
if(variance){
  ret$var.seq=est.var.seq
  ret$inf.func=inf.est.func
}

class(ret) <- c("crrQR")
return(ret)
```

```
}
```