Survival of the Fit and the Fat

Decomposing the exit hazard of Norwegian firms

Anja Grinde Myrann

Master Thesis for Master of Philosophy in Economics

Department of Economics

UNIVERSITETET I OSLO

February 2010

Preface

This master thesis was written at the Ragnar Frisch Centre for Economic Research in Oslo. It is part of the project *R&D, Industry Dynamics and Public Policy*, Frisch project no. 6512. The project is funded by the Norwegian Research Council, and is a joint project with researchers from the Department of Economics at the University of Oslo, the BI Norwegian School of Management and the Frisch Centre.

I wish to thank my supervisor Knut Røed, for giving me the opportunity to work with these interesting and challenging topics and methods, and for providing valuable ideas and comments throughout the process. Simen Gaure deserves special thanks for making the estimation of the models possible. I am also grateful to all my colleagues at the Frisch Centre for answering my questions, providing comments, and being a joy to spend time with. Lastly, I want to thank Ole Røgeberg, for comments and proof-reading as well as for his ceaseless support, encouragement and patience.

Oslo, February 1st 2010

Anja Grinde Myrann

Summary

This paper considers the fate of Norwegian firms in their first decade after entry. The underlying dynamics of entry and exit play an important part in the growth and development of an economy. On the one hand, there is net entry of firms in some industries or sectors and down-sizing in others to adapt to changing economic realities, such as demand shifts and relocation of production. On the other hand, there is a considerable excess turnover of firms within industries, which some researchers have explained as productivity-based sorting. This means that the entry and exit of different firms are intrinsically linked. The survival and growth of some comes at the expense of others.

This paper will consider the following problems: 1) What is the duration dependence of the exit hazard of an individual firm? Is the duration dependence positive or negative? 2) Can we observe any differences between the duration dependence of different types of exits? 3) What factors other than age can contribute to an explanation of why some firms exit while other survive? Are there differences in how these factors impact the risk of different types of exit?

Economic theory provides several different mechanisms that may account for differences in firm exit rates. Vintage theories emphasize the age of capital and rate of technological innovation. These are also known as theories of creative destruction or Schumpeterian growth. They predict an increasing exit hazard in the age of capital. Theories of learning can be divided into passive “learning about ones relative quality through market feedback” and active “learning-by-doing” (improvement in quality with time and experience). Passive learning is treated explicitly in this analysis, by both taking into account the observed differences and modeling the unobserved differences that exist between firms. These differences mean that firms are of different quality and have differing exit risks. This creates a sorting process. Since "low quality" firms tend to exit earlier than higher quality firms, we get a selection process that makes the “gross” observed exit hazard (i.e. ignoring differences in quality) decline in firm age. In addition to the two groups of theories mentioned above, business cycles and sector shifts have effects on the entry and exit of firms that differ by year and by industry/sector. These effects are not the focus of my analysis, but are controlled for by year and sector dummies.

The analysis has a particularly rich categorization of the different types of entries and exits. Using a comprehensive set of register data, we are able to distinguish between the plant and the firm level. This allows us to control for different types of entries, and identify those cases where the plant survives an exiting firm. We are also able to distinguish bankruptcies from other exits. The identification of different types of firm exit is exploited in two models of competing risks. Using a model with competing risks, we are also able to identify differences in the factors associated with different types of exit.

To decompose the gross exit rate I use a series of proportional hazard models with parametric and non-parametric assumptions regarding the unobserved heterogeneity. One of the models assumes that the unobserved heterogeneity in the sample is gamma-distributed. This model is estimated in a Stata program developed by Jenkins (1997). The other models are non-parametric in that they estimate a discrete distribution of the unobserved heterogeneity of firm. These models are estimated by an R-program developed by Simen Gaure at the Frisch Centre. By explicitly modeling the unobserved quality differences between the firms, we are able to separate out the selection/sorting effect from the duration dependency facing individual firms. The results turn out to be similar for the model with a gamma-distributed unobserved heterogeneity and the non-parametric model.

The analysis is performed in a number of stages. The observed hazard rate of firms is declining in firm age. . Taking differences in the observed heterogeneity of firms into account, the remaining duration dependency of the exit hazard is not significantly related to firm age. Estimating a model with unobserved heterogeneity turns the duration dependence strictly positive, which indicates that the exit hazard of firms increases with firm age when selection effects are taken into account.

Introducing competing risk models, I distinguish first between the cases of firm exits where the corresponding plant also shuts down ("full exit") and the cases where the plant continues under a new firm ("half exit"). I find that the duration dependency of half exits is more positive than that of full exits, meaning that a continuation of the plant after firm exit is relatively more common if the firm is older at time of exit. In the second competing risks model, bankruptcy is distinguished from other firm liquidations. The results indicate that bankruptcies differ significantly from other liquidations. First, the bankruptcy hazard has no discernible duration dependency, our results are not significantly different from one where the risk of bankruptcy is constant with firm age given firm quality. This suggests that the positive duration dependency of the total exit hazard is driven solely by the non-bankrupt exits. Second, the observed differences between firms, such as employment and debt-to-equity in the year of entry, are associated in starkly differing ways with patterns of bankruptcy hazard compared to liquidation hazard. For example, a firm with negative ratio of debt-to-equity has a substantially higher risk of bankruptcy than a firm with moderate debt-to-equity, but seems to have lower risk of a non-bankrupt exit.

Some of the results were significant in interesting ways. Interestingly, a high share of female employees in the year of entry is associated with a lower risk of bankruptcy. The results in the single risk model indicates no relationship between the initial employment size and the exit hazard. This surprising finding can be explained in the competing risks model, where the employment size has opposite effects on bankrupt and non-bankrupt exits. High employment is associated with a reduced risk of bankruptcy and a higher risk of liquidation. The competing risk models demonstrates that ignoring the distinctions between types of exits obscures underlying differences in the duration dependencies for different types of exits.

Contents

List of figures and tables:

1 Introduction

The entry and exit of different firms are intrinsically linked. The survival and growth of some comes at the expense of others, and the underlying dynamics of entry and exit play an important part in the growth and development of an economy. On the one hand, there is net entry of firms in some industries or sectors and down-sizing in others to adapt to changing economic realities, such as demand shifts and relocation of production. On the other hand, there is a considerable excess turnover of firms within industries, which some researchers have explained as productivity-based sorting. Huse (2009), for instance, has shown that the entry and exit of firms contributes significantly to the overall productivity growth in the economy, especially in times of economic slow-down, as low-productivity firms exit while firms with higher productivity enter the marked. Foster et al. (2006) finds that net entry of firms accounts for the majority of the changes in labour productivity in the U.S. retail trade sector. An overview of the research on productivity and turnover is found in Caves (1998).

The productivity enhancing effect of entry and exit is a combination of factors within and between firms. I will discuss this with reference to the differing *quality* of firms, quality here referring to the post-entry performance of firms in terms of productivity or profitability. On the one hand, a firm can be forced to exit because it has a permanently lower quality than the other firms in the market. This is the process of selection, the quality firms survive while the low quality firms exit. The selection process can also be referred to as passive learning. On the other hand, when new firms with increasingly higher quality then the older firms enters the market, the old firms that were once the best of their kind will be forced to exit. This process is discussed in vintage theories of firm entry and exit.

Theories on firm entry and exit can be broadly grouped in two branches, vintage theories and learning theories (Dale-Olsen, 2005). The vintage theories are also known as theories of creative destruction or Schumpeterian growth. They are based on the notion that new productive capital is required in order to make use of technological innovations. Old capital is thus less productive than new capital, and more so the higher the rate of technological progress. Old capital is thus a proxy for low efficiency (Salvanes and Tveteras, 2004). For each new step in the process of technological innovation, some old capital with old technology will be rendered unprofitable by the entry of new, more productive methods. A firm with old unprofitable capital will either exit, or invest in a renewal of capital. The vintage theories predict an increasing exit hazard in the age of capital. Testing this theory could be based on the assumption of a link between firm age and capital age. Dale-Olsen (2005) finds support for the vintage theories when modelling the exit hazard of plants based on data for a period of forty years. He finds that the hazard function is significantly increasing in plant age. Salvanes and Tveteras (2004) present an analysis where the distinction between the age of capital and firm age is made explicit. They have access to detailed investment data which makes it possible to compute a measure of capital age based on past investments. The estimation results strongly support a separation of capital age and plant age, as the learning and capital replacement effects work in opposite directions with respect to exit probability. The likelihood of plant shutdown is significantly decreasing and convex in plant age and significantly increasing in the age of machine capital (Salvanes and Tveteras, 2004).

The second group of theories on entry and exit contains theories of learning or experience. When a new firm enters the marked, two learning processes are commenced, referred to as passive and active learning. In passive learning, the firm learns about its own profitability, based on both its own quality and the actual economic circumstances it faces after entry. Some potential entrepreneurs have ideas that they decide not to carry out, because the business prospects are too risky or because the expected profitability is not sufficiently high. These "unborn firms" are obviously never observed. Other ideas are considered good enough to be given a chance. Still, unresolved uncertainties remain which cannot be known before entry. In the first few years after entry, the entering firms learn whether they are viable or not. If the expected future profits are too low, the firms exit. This creates a selection process.

The firms with low profitability, the "bad" firms, will have a higher probability of exit at any given point. On average, they will also exit earlier than the "good" firms. The good firms will therefore constitute a rising share of the population of remaining firms as they get older. Thus, if the selection process of passive learning was the only force affecting the duration dependence of firm survival, the exit rates would be decreasing in age. Jovanovic (1982) illustrates the selection process through a model of *noisy selection*, in which the cost function is not only unknown to the firm before entry, it also differs from firm to firm. Dale-Olsen (2005) follow the cohort of plants established in 1996 during their first seven years. The exit probability for small firms, with maximum ten employees, falls from more than 20 percent the first year to eight percent for the seventh year. The exit hazard for small firms is rapidly declining in age, at least until the sixth year, but for the larger firms there is no clear trend (ibid).

The gross hazard rate is the average exit hazard of the remaining firms in the sample at any given age. This rate seems to fall over a firm's first few years, which is often interpreted as the selection effect dominating the gross hazard during the early years of a plant's life-cycle. Though passive learning takes place within individual firms, the effect on the exit hazard over time is at the firm population level. It pulls the observed gross exit rate in the direction of declining hazard as firm age increases. But the passive learning does not alter the hazard facing an individual firm, except in that the firm can be increasingly confident that it has high quality as it ages, because age is depending on survival up to the present.

The other learning process is active learning within each firm. The idea is that the firm quality is gradually improved over time by experience and increasing knowledge. The firm learns about its economic environment and the business it has entered, and adapts by improving decision making, management and production processes. Both active and passive learning causes the expected exit hazard of any individual firm to decrease in age. Passive learning has this effect because the expected exit hazard increases when we know the firm has already survived a number of years, survival indicates that the firm is of higher quality than the average entering firm. Active learning has a direct effect on the quality of the individual firm, and as the quality increases, the exit hazard decreases.

The theories can be seen as complementary theories describing different processes and thus different forces at work. They may be applicable to different time periods, different parts of a firm's life-cycle, and they may depend to varying degrees on factors such as industry characteristics, speed of innovation and technology. In a capital intensive sector with rapid technological growth the effect of creative destruction is likely to kick in at an earlier stage than in a labour intensive, static industry. Likewise, learning by doing can be expected to play a larger role for a longer time for a firm entering an industry with an advanced, knowledge intensive technology.

This paper will focus on the exit hazard of new firms. The risk of exit varies both across firms and within firms over their life-cycle. I will attempt to decompose the gross exit hazard observed in the population of firms, separating the selection process from the duration dependence of the hazard rate. The duration dependence is the net outcome of a number of mechanisms, amongst them the vintage and the active learning effects, which both operate at the firm level. However, a number of other factors are likely to influence the duration dependence, though these are not usually formally discussed in relation to theories of entry and exit. For example, a dentist running a private practise will often let his firm be liquidated when he retires. A restaurant or pub owner might want to start afresh with a new firm operating in the same location if the guests seem to get bored with the old concept. Other firms are liquidated because the owner want to move on, get a new job, start a new business or simply cash out the values in the firm. Considering these alternative motivations for firm exit, the picture of what can be expected about the relation between firm age and the probability of exit becomes more complex.

The selection effect, as mentioned, works at the population level, and is caused by differences between firms. To the extent that these differences are observable, these differences can be controlled for in an analysis of duration dependence. But limitations in our data will always leave unobserved heterogeneity; the true quality of an individual firm at time of entry is unobservable. The key to overcoming this problem and get past the effect of selection is to explicitly take unobserved heterogeneity into account when estimating duration dependence.

As an extension to the decomposition of exit rates, this paper will also consider whether the duration dependence is different for different types of exit. The data allows for a separation of different types of firm exits along two dimensions. First, some firm exits are characterized by simultaneous plant exits, while in other cases the plant continues under a new firm. Second, firms exits due to bankruptcy are separated from non-bankrupt firm exits.

To summarize the key questions I will attempt to answer:

1) What is the duration dependence of the exit hazard of an individual firm? Is the duration dependence positive or negative?

2) Can we observe any differences between the duration dependence of different types of exits?

3) What factors other than age can contribute to an explanation of why some firms exit while other survive? Are there differences in how these factors impact the risk of different types of exit?

2 The ideal experiment and an executable strategy

My analysis is based on four different data sets, all originating from Norwegian administrative registers. They contain employment data, accounts data, demographic data and firm data from the period 1995 to 2005. An overview of the four data sets is found in Table 1.

I want to look at firm entries and assess which firms turn out successful and which firms exit. Ideally, I would have had sufficient information to perfectly characterize the initial and constant parameters that are relevant for firm quality and survival. Because I can not know all the relevant characteristics of the firms in the sample, an important aspect of this paper is to find methods to deal with the unobserved differences between firms.

Ideally, I would have data on all the Norwegian firms for prolonged period of time. The availability of data, however, imposes extensive limitations on the firm population I can use in the analysis. On the one hand, I only have data from eleven years, and need two of them for defining entry and exit. That leaves me with a maximum of nine observation years per firm. On the other hand, there is a trade-off between the wish for a large sample and the wish for detailed information about the firms included in the sample.

At the very core of the notion of a selection effect lies the fact that firms are different. The larger the heterogeneity in the firm population, the larger are the consequences of limiting the sample if the purpose is to extrapolate the results to the population of firms. For instance, missing values for a variable in a register, such as county, industry or sex and age of CEO, was found to have a significant positive effect on the estimated exit hazard, meaning that the exit hazard is higher for firms with incomplete register data. Thus missing information seems to be a bad sign for firm survival. By excluding some firms for which the availability of data is insufficient, I am consequently left with a population of higher-quality firms in the analysis.

Preferably, the data should also allow a separation of capital age and firm age, by containing full information on the age of the capital and the technology used in the production processes of each firm. As this is not the case, I will not attempt to separate vintage effects from active learning in the estimation of duration dependence. The assessment of capital and labour input will be discussed later.

Lastly, structural changes within and between industries and the market structure in which the firm operate is likely to influence the decisions of entry and exit. I use control variables for industry and geographical location in order to capture some of the effects from the market environment of each firm.

Data set	Description	Years	Data type	Level	Source	Limitations
Employment data	Details on each employer-employee relationship	1995-2005	panel	Firm, plant, individual	Statistics Norway	Does not include self-employment
Demographic data	Individual characteristics such as age, education and sex	1995-2005	panel	Individual	Statistics Norway	
Accounts data	Annual accounts of Norwegian firms	1995-2005	panel	Firm	Brønnøysund Register Centre	Not all firms have an obligation to report annual accounts
Firm data	Firm characteristics such as industry, county, registration and bankruptcy/liquidation dummies	extracted in 2007	cross-section	Firm	Brønnøysund Register Centre	

Table 1 - Data sources

3 The models

The starting point for my analysis is the observed, "gross" exit rate in my sample, depicted in Figure 1. It peaks at nine percent after three years and then gradually declines to about six percent. The first model will control for year effects and known differences between firms, such as industry, county, size and characteristics of the employees. I proceed to models which explicitly account for unobserved heterogeneity of firms in order to disentangle the selection effect from the duration dependence. At last, I distinguish between different types of exits in a model with competing risks.

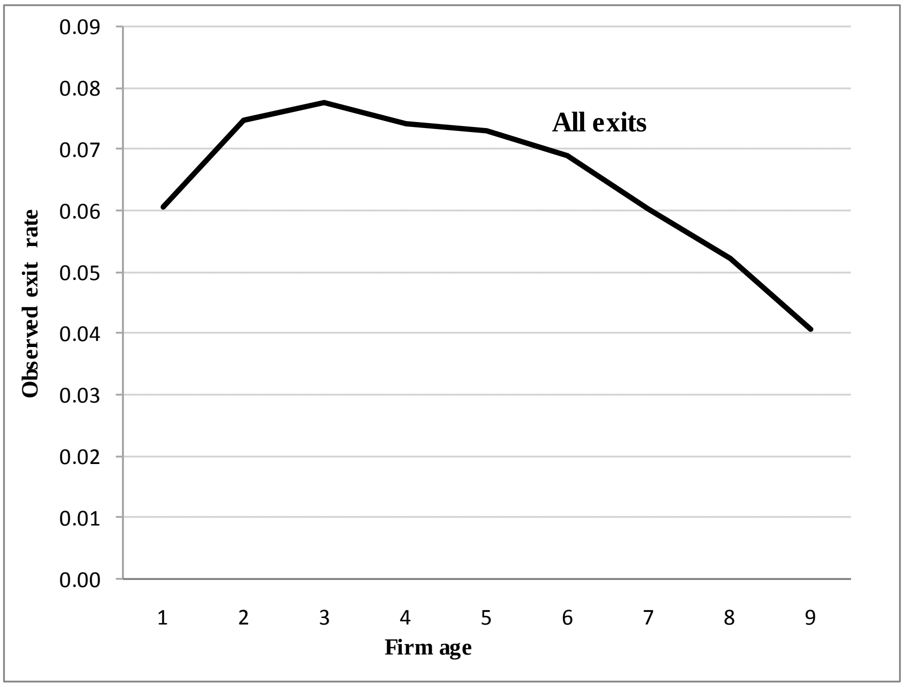

Figure 1 - Observed exit rate, percent

3.1 The proportional hazards model

The description of the model set-up is based on Jenkins (1997) and Røed and Westlie (2007). I assume that the hazard rate function for firm i at time $t > 0$ takes the proportional hazards form

$$\lambda_{it} = \lambda_0(t) \cdot \exp(\mathbf{x}_{it}'\beta),$$

where $\lambda_0(t)$ is the baseline hazard function. In a proportional hazard model, the effect of covariates can only induce proportional shifts in the hazard rate but can not change its shape (Blossfeld et al., 2007). Put differently, the effect of a time-invariant variable is to scale the risk profile of a firm up or down proportionally at all points - if it raises the risk at some point in time by 10%, it also raises the risk at all other times by 10%. It can not, however, create a shift which increases the risk of exit uniformly e.g. with two percentage points at all points in time.

The discrete time representation of the continuous proportional hazard model is also known as a complementary log-logistic model (clog-log). While other models, such as the Weibull model, force a smoothly increasing *or* decreasing rate onto the results, the log-logistic model allows for a non-monotonic hazard rate. This means that the exit probability in the clog-log model used here can be increasing for some periods of time and decreasing in others. On discrete form, the time hazard in the jth interval lasting from time a_{j-1} to a_j can be written as

$$h_j(\mathbf{x}_{ij}) = 1 - \exp\left[-\exp(\mathbf{x}_{ij}'\beta + \gamma_j)\right] \text{ with } \gamma_j = \log \int_{a_{j-1}}^{a_j} \lambda_0(\tau)d\tau,$$

where γ_j represents the duration dependence of the hazard, also called the baseline hazard. With a non-parametric baseline hazard with a separate parameter for each duration interval, γ_j is interpreted as the integral of the baseline hazard over the relevant time interval. It gives us the factor by which the hazard function is scaled for each year compared to the reference year.

3.2 Unobserved heterogeneity

Distinguishing between the effects of type and the effects of time is a main focus in this paper. If there is unobserved heterogeneity or omitted variables, the time hazard estimated in the basic model will include a selection effect. The idea is that, even after controlling for all the observable differences between firms, unobserved quality differences remain, giving some firms a higher hazard than other firms. Imagine that half the firms starting up have zero probability of failure, while the other half have a 50% chance of failure in any given year. The first year, 25% of all the new firms fail. The next year, the bad firms are only a third of the remaining firms - so when half of these fail in the second year the overall observed failure rate has fallen from 25% in the first year to 1/6 in the second. The next year, the remaining bad firms are only 20% of the surviving population, bringing that year's observed failure rate to 10%. Every year the failure rate falls - but this is in our example purely a result of permanent and fixed quality differences between firms - it's pure selection. Returning to the real world of firm exit, there are differences between the firms in our sample that we can not observe. If there are differences in the exit probabilities of firms, then as time goes by the low-risk firms will make up an increasing share of the firms remaining in the sample. The average exit hazard therefore declines, as we could see in Figure 1.

The presence of unobserved heterogeneity is not unique for event history analysis. But unlike for example a cross-section regression, the unobserved heterogeneity in event history analysis is not evenly distributed across spell durations. From the outset, ν is white noise, but as years go by, it causes a sorting of firms according to their unobserved differences in quality. This results in a spurious negative duration hazard.

In order to account for the presence of unobserved heterogeneity, the basic model setup is extended by adding a firm specific unobserved covariate ε_i. I then get a mixed proportional hazard model where the instantaneous hazard rate is

$$\lambda_{it} = \lambda_0(t) \cdot \varepsilon_i \cdot \exp(\mathbf{x}_{it}'\beta)$$

with the corresponding discrete time hazard function

$$h_j(\mathbf{x}_{ij}) = 1 - \exp\left\{-\exp\left[\mathbf{x}_{ij}'\beta + \gamma_j + \nu_i\right]\right\} \text{ where } \nu_i = \log(\varepsilon_i).$$

In order to estimate this model, some further assumptions on the distribution of ν must be made. A Gamma-distribution is often used in the literature. The Gamma-distribution has become very popular, at least partially owing to the fact that it simplifies the calculations of the model, but also the normal distribution and other are possible. Abbring and Van den Berg (2007) presents a rationalization of the preference for the Gamma-distribution in duration analysis. However, imposing a Gamma distribution on the unobserved heterogeneity, or for that matter any other given distribution, has consequences for the estimated hazard rate. If the true distribution of the unobserved heterogeneity in the sample is far from the Gamma distribution, imposing such a distribution on the model can bias the estimated duration dependence. Because there is no evidence to support the assumption of a Gamma-distribution, I prefer to use a less restrictive model which approximates a discrete distribution of the unobserved heterogeneity.

Nevertheless, I want to check different assumptions about the distribution of unobserved heterogeneity and compare it to the model with a discrete distribution. The basic model is the reference model, which makes no attempt to filter out the selection effect. I also estimate one model which follows the common assumption of Gamma-distributed heterogeneity. In this model, ν is a Gamma-distributed random variable with unit mean and variance σ^2. The model with Gamma-distributed heterogeneity consequently has one extra parameter compared to the basic model, namely the variance of ν (Jenkins, 1997).

Approximating the unknown distribution of unobserved heterogeneity by means of a discrete distribution is modelled by means of a non-parametric maximum likelihood estimator (NPMLE). The model estimation proceeds as follows. First a null-model disregarding heterogeneity is estimated. This corresponds to the basic model above. Second, the unobserved heterogeneity of each firm is allowed to take two different values ν_1 and ν_2. Two new parameters are estimated in this step, the two values for ν and the probability $p = p_1$ that a firm belongs to group 1. Then the number of different values of the unobserved heterogeneity ν is expanded step-wise. The iteration continues until there are no further gains from expansion according to the maximum likelihood principle. The maximum likelihood principle of estimation is based on the idea that the sample of data at hand is more likely to come from a "real world" in which the parameter values are the maximum likelihood parameters than from a world with any other set of parameter values (Kennedy, 2008; Greene, 2003).

Gaure et al. (2007) has conducted an extensive Monte Carlo assessment of the non-parametric maximum likelihood estimator. They conclude that it is extremely reliable, provided that the sample size is large and that there is some exogenous variation in the hazard rates (Gaure et al., 2007). The method is very robust to differences in the underlying distribution of unobserved heterogeneity, and right-censoring of the sample data is not problematic. However, serious bias problems can arise if the assumption of mixed proportional hazard does not hold. Gaure et al. (2007) has shown that the NPMLE method will approach a distribution which is very close to the actual, underlying distribution of unobserved heterogeneity in the sample.

The method with discrete modelling of the unobserved heterogeneity in the sample facilitates a separation of the effect of firms being different from the effect of firm age. The effect of firms being different is captured by the differences in their unobserved heterogeneity, given by the different ν's reported, and the assigned probabilities for each level of ν. The effect of firm age is the remaining part of the duration dependence once the heterogeneity has been filtered out. However, the estimated support points for the distribution of $\nu$and their assigned probabilities can not be interpreted directly. Rather, they are a means to reaching estimates of duration dependence without noise from the selection process.

3.3 Competing risks

The data used in the analysis has information that allows a distinction between different types of exit. I will separate the firm exits along two dimensions; (1) whether or not the plant is still operating after the firm exits, and (2) whether the firm is bankrupt or liquidated. In order to investigate whether the duration dependence of different types of exit have similar profiles or not, I will use a transition model with two competing risks. The two hazards are dependent and must therefore be modelled simultaneously. When modelling two competing risks, I assume a piecewise constant hazard (Røed and Westlie, 2007). In other words, the hazard is constant within each period, so that $\lambda_0(t)=\lambda_j$ for $\tau-1<t<\tau$ and thus $\gamma_j=\lambda_j$. In this setting, I no longer need to distinguish between period and time, so I will drop the notation with j for period and use t to indicate period.

There are two baseline hazards in the competing risks model, λ_{1t} and λ_{2t}. The competing risks models are also estimated with the NPMLE, modelling the unobserved heterogeneity by means of a discrete approximation. In the competing risk-setting, ν_i becomes a vector with one value for each transition, ν_{1i} and ν_{2i}. All parameters in the model are estimated separately for the two transitions. In order to simplify notation, I define an index function $w_{kit}=\beta_k x_{it}+\lambda_{kd}d_{it}$ for $k=1,2$, where x_{it} is a vector of observed firm characteristics and calendar dummies and d_{it} is the vector of period dummies. With two possible transition, the probability of transition of type 1 in period t for firm i is

$$h_1(w_{1it}+\nu_{1i})=\left\{1-\exp\left[-\exp(w_{1it}+\nu_{1i})-\exp(w_{2it}+\nu_{2i})\right]\right\}\frac{\exp(w_{1it}+\nu_{1i})}{\exp(w_{1it}+\nu_{1i})+\exp(w_{2it}+\nu_{2i})}.$$

The first parenthesis gives us the probability of exit, and the fraction is the probability that the exit is of type 1.

The likelihood function for the competing risks model in the NPMLE-framework is derived in Røed and Westlie (2007). Let K_{it} be the set of feasible transitions for firm i in period t, which in the model with two competing risks is limited to $K_{it} = 1,2$. Let y_{kit} be an outcome indicator, equal to one if firm i undergoes transition k in period t and zero otherwise, and let Y_i be the complete set of outcome indicators available for firm i, in this case maximum nine periods per firm. The contribution to the likelihood function from an individual firm is then given by

$$L_i(\nu_i) = \prod_{y_{kit} \in Y_i} \left[\prod_{k \in K_{it}} \left[\left[\left(1 - \exp\left(-\sum_{k \in K_{it}} \exp(w_{kit} + \nu_{ki}) \right) \right) \frac{\exp(w_{kit} + \nu_{ki})}{\sum_{k \in K_{it}} \exp(w_{kit} + \nu_{ki})} \right]^{y_{kit}} \right] \times \left[\exp\left(-\sum_{k \in K_{it}} \exp(w_{kit} + \nu_{ki}) \right) \right]^{\left(1 - \sum_{k \in K_{it}} y_{kit}\right)} \right].$$

In the single risk model presented in the previous section, where $K_{it} = 1$, this expression can be simplified to

$$L_i(\nu_i) = \prod_{y_{it} \in Y_i} \left\{ \left[\left[\left(1 - \exp\left(-\exp(w_{it} + \nu_i) \right) \right) \right]^{y_{it}} \right] \times \left[\left[\exp\left(-\exp(w_{it} + \nu_i) \right) \right]^{(1 - y_{it})} \right] \right\}.$$

4 The data set

This chapter will begin with an identification of my sample, and the definitions of entry and exit on which the analysis is based. Next, I explain the source and definitions of the variables used as explanatory variables and controls in the analysis.

4.1 Sample selection, entry and exit

When defining whether a firm should be included in the sample, I have used criterions both regarding the availability of data, the type of firm, and its process of entry. A summary of the steps from the merged data set with all the firm level information to the selected sample is found in Table 2.

		Firms	Firm-years
1	All observations	399885	2565420
2	Not present in 1995	215087	1031781
3	In *accounts data*	157366	723256
4	Ltd. firm in first year of accounts, year T	130662	629934
5	In *employment data*	72502	411234
6	Employment in year T	49343	268820
7	No employment before year T-1	49327	268730
8	Singleplant firm in year T	48076	240620
9	Ltd. singleplant firm first year in employment data (if that is year T-1)	48069	240566
10	In *firm data*	47972	240438
11	Registration and founding completed no earlier than year T-3	45460	227092
Final sample	Remaining sample after excluding 2005-entries and right-censoring no-exit firms	**41122**	**177105**

Table 2 - Sample selection

4.1.1 The level of analysis: Firm vs. plant

The employment data and the demographic data can be liked by individual identification numbers. The employment data describes all employer-employee relationships, and covers the whole labour marked in Norway except for the self-employed. Each entry has variables identifying three different levels, the individual employee, the plant and the firm.

The availability of both firm- and plant-level organizational numbers in the same register is of great value when I want to identify and characterize firm entry and exit. In order to clarify the distinction between the two levels, I will give a brief definition of the concepts as used in this paper. The *firm* is the smallest combination of judicial units to constitute an organizational unit producing goods and services. The word *plant* refers to the location of economic activity, regardless of industry (SSB, 2010a). A grocer's shop, a pharmacy, a shipyard, a restaurant, a tailor and a news stall are all unique plants in this respect.

The firm is the judicial owner of the plant. One firm might own more than one plant, but most often they stand in a one-to-one relationship, one firm owning one plant[1]. I will then refer to it as a *single-plant firm*. Multi-plant firms can operate in several industries and geographical locations simultaneously, and would therefore be difficult to characterize along these dimensions. This is the reason why only single plant firms will be included in my analysis of entrant firms.

On the other hand, a firm may itself be part of a group of firms. Such a group will report consolidated accounts, in addition to the accounts of each individual firm in the group. In the consolidated accounts the group is accounted as one unit, so all transactions between the firms inside the group will be disregarded. Such consolidated accounts are distinguishable in the accounts data, and thus they have been excluded to avoid double recording of the same economic activity. Unfortunately, I can not tell which firms belong to such a group and which are independent, so some of the firms in my analysis might be part of a group.

4.1.2 Limited liability firms and the obligation to report accounts

The Norwegian accounting act (Regnskapsloven, 1998) establishes the types of firms that are legally obliged to keep accounts and report them to administrative registers. All firms organized as limited liability (Ltd.) firms, both private and public[2], have such an obligation. The accounts data contain all the reported annual accounts of Norwegian firms from 1995 to 2005. The availability of accounts data from limited liability firms is the reason why this paper will only follow Ltd. firms. The accounts are on the firm level, so the firm will be the unit of analysis.

[1] 95% of the firms in the employment data from 1995 to 2005 are single plant firms.
[2] Norwegian organizational forms *AS* and *ASA*, *aksjeselskap* and *allmennaksjeselskap*

On average for the years 1996 to 2004, the Ltd. firms comprise 95 percent of the firms in the accounts data, 61 percent of firms in the employment data and 56 percent of total employment, as calculated from the employment data. Ltd. firms made up about 55 percent of all firms and 52 percent of total employment in 1996, increasing to 66 percent of the firms and 57 percent of employment in 2004.

Inclusion in the sample require that the firm is single plant limited liability firm at time of entry. If, however, the firm later changes organizational form or becomes a multi-plant firm, it will be right-censored from that point in time.

4.1.3 Separating the stories of firm and plant

When using this type of register data, it is not possible to follow the life-cycle of a firm directly, as the firm is only observed through its organizational number, and the organizational numbers attached to a firm or a plant can change from one period to the next for a number of different reasons. A plant's firm number may change if the firm changes its organizational form, say, from sole proprietorship to private limited liability company, or if the plant is sold to another firm. The plant number can change if the plant is moved to a new location. By considering data on both levels, however, it becomes possible to record a more detailed story than would have been possible with either firm- or plant-level data alone. I separate the new firms – the new organizational numbers on the firm level – into subgroups depending on whether the plant is also brand new or has a history in advance of the firm entry. If the plant is older than the firm, the characteristics of the previous firm attached to the plant in question is known. What was its organizational form? Was it a single plant firm? The same separation is done at the other end of the firm's life-cycle. When a firm number exits from our register, we can check whether the plant continues or not, and, if it continues, then in what type of firm. For example, many firms begin as sole proprietorships and later change organizational form to become limited companies. They then become *new firms*, but we will obviously expect their characteristics and their hazards to be different from those of *genuine start-ups* – where the plant is also brand new. By paying attention to both levels of the firm-plant structure, it is possible to distinguish between exits due to failure and more successful exits such as takeovers.

4.1.4 The availability of data

Combining information from the different registers imposes restrictions on the observations that can be included. Only firms present in all data sources were included in the analysis. This implies that I only considered firms with (a) reported accounts, (b) employees and (c) available firm data.

There are further difficulties related to the inconsistent availability of data in the employment and accounts register for each operating year of the firm. The selection of Ltd. firms only is motivated by their common obligation to report accounts. This legal obligation implies that any Ltd. firm found in the employment register should be found in the accounts data for the same year. This is not always the case. Most disturbingly, many firms have paid employees for a prolonged period before they first report accounts. Also, many firms don’t report accounts for their last couple of years with paid employees. The treatment of firm-years with no accounts or no employment will be described in the next section, when defining entry and exit.

4.1.5 A restrictive definition of entry

Defining entry requires some trade-off between the desire to obtain a large sample and avoiding the risk of false entries, that is, defining as new a firm which has already been in operation for a long period of time.

Salvanes and Tveteras (2004) have compared rules for identifying entry and exit for a data set similar to this one, covering Norwegian manufacturing firms in the period 1977-1992. They have access to explicit identification of new plants, continuing plants, and plants that closed down in the years 1977-1986. The second way of identifying entry and exit is to define as entering a firm that appears in the date in year t without being observed in year t-1, and likewise, define as exiting any firm present in year t but not in year t+1. After comparing the two identification strategies, they find only negligible differences in the number of exit. From lack of explicit information on entry and exit, I will therefore use the latter procedure, although slightly altered to exclude some kinds of false entries.

I use three different data sources to define entry and exit.

1. Reported accounts: In the sample of single plant Ltd. firms, a firm is defined as entering in year t if it has reported accounts in this year but not in year t-1. This implies that the first year of entry in the sample is 1996, observations from 1995 are used to define entry.

2. Employment records: A new entering firm must have positive employment in the year of entry, that is, it must be in the employment register for year t. If I had included a firm which had reported accounts one year before first employment, the firm's inclusion would have been dependent on survival through the first year and would bias the sample. The employment requirement excludes the Ltd. firms that are only established to store assets or for other non-productive purposes. Unfortunately, this also excludes some firms that might be interesting but that start up slowly or at a very small scale. Omitting these was judged to be better than including the above mentioned "empty" firms.

If the firm had employees as early as in year t-2, it will also be excluded. As explained in the previous section, any Ltd. firm with employees but missing accounts data have failed to fulfil their legal obligation to report accounts. Some errors in dating of the observations makes one year of employment before first accounts (employment in year t-1) acceptable, but the observation from this year will be excluded from the analysis.

3. Firm data: Lastly, I require that the founding of the firm and official registration was completed no earlier than in year t-3.

Hence, firms with long lags between the time of registration, first employment and first accounts, will be omitted from the analysis. Excluding them reduces the likelihood of false entries in the analysis, but it also limits the sample of firms. Some firms will never get defined as entering according to the above definition.

When running the regression I will use dummies to distinguishing five different transitions at time of entry. The most common transition is the full entry, where the plant enters at the same time as the firm. The remaining four dummies describe the opposite cases, when the plant had a history before firm entry. Each of the four dummies represent different characteristics of the firm which the plant previously belonged to, grouped according to whether or not it was a Ltd. firm and whether it was a single-plant or a multi-plant firm.

4.1.6 Defining and characterizing exit

Several sources of information contribute to the definition and characterization of firm exit. Observations of the firm's last year of reported accounts and the last year of employment is supplemented with the firm data, which includes dummy variables with information from 2007 on whether the firm is still active or deleted and whether it has entered a bankruptcy process or not.

After excluding the firms with several years of employment before the first reported accounts through the definition of entry, many firm-years remain that lack reported accounts in the other end of their life-cycle. When considering all observations after entry for the firms in the sample, there are 7272 firm-years with employment but no accounts, constituting 3.4% of the observation-years in the sample. All of these years are after the last reported accounts, and firms that end in bankruptcy are responsible for 83% of these firm-years with missing accounts. It seems that firms undergoing a bankruptcy process do not give priority to reporting the usual annual accounts, they rather spend their resources elsewhere. Only one third of the firms who go bankrupt report accounts for the last year of employment, compared to 97% for the rest of the exiting firms. The data suggests that the firm has usually already entered the bankruptcy process, or is speeding rapidly towards it, if it has employment but do not report accounts. Because it appears that not reporting accounts is a clear sign of imminent final exit, I disregard the employment which take place after the last year of accounts. The firm is defined as exiting in year t if it has reported accounts for year t but not for year t+1. The last year of exit is 2004, firms who are still present in the accounts data in 2005 have unknown outcome, and is right-censored after 2004.

There are also many cases where employment ceases well before the firm is formally closed down and deleted, and there can be years of apparent non-activity, in which the firm has a temporary halt in employment. Continuous employment until exit is not required and the employment level after the first year will not be considered in the analysis.

The data allows for a separation of different types of firm exit in two respects. First, as explained in section 4.1.3, it is possible to distinguish between the exits where the plant continues under another firm and simultaneous exit of firm and plant. Second, the firm data has information on bankruptcies, which makes it possible to separate bankruptcies from other firm exits. Non-bankrupt exits will be referred to as liquidations.

4.1.7 Summing up entry and exit

In total, there are 41 122 unique firms in the sample, with 177 105 observation years altogether. 12 235 firms are defined as exiting, constituting nearly 30% of the firms we follow. The average exit rate is 6.9% per firm year.

Through the above definitions, the birth and death of the firm follows the firm's presence in the accounts data. Thus, all of the firm-years with missing accounts are excluded. This increases the reliability of our definition of entry. Observing that none of the firms in the sample has an intermediate year where accounts are not reported, it is quite unlikely that any of the firms defined as entering actually did exist before 1995.

The information in the firm data confirms that nearly 99 percent of the defined exits are indeed listed as bankrupt or otherwise deleted by 2007. Of the remaining 166 unconfirmed exits, there are only nine exits before 2003. The lack of data in the last two years indicates that there are some lags before registration rather than extensive false exits.

The data covers the years 1995-2005. The first and last years are used for identifying entry and exit, so they can not be included in the analysis. Any firm present in the accounts data or employment data in 1995 will be excluded, so the first observed entry of firms is in 1996. Any firm still in the accounts data in 2005 is a no-exit firm, while a firm present in 2004 but not in 2005 is defined as exiting in 2004. That leaves me with a maximum of nine observation-years per firm. There are more observations for the early years of a firm's life-cycle than for later years, because the firms entering late in the period 1996-2004 will be observed for a shorter period of time.

The definitions of entry and exit have some impact on the resulting estimates of the exit hazard's duration dependence. Defining the first year in any data set as entry, or the last year in any set as exit, would stretch the hazard curve horizontally, but it is not likely that it would significantly alter its general shape. Excluding all firms who never report accounts, on the other hand, could influence my analysis. Some firms exit before they even get as far as to report their first annual account. There are about 3000 single plant Ltd. firms that are observed with employees for a few years but never report accounts, 67 percent of these last for only one year and 94 percent exit before their third year. Their absence in my selection of firms will cause the exit rate for the firms in my sample to be lower than the true exit rate for the population of Ltd. firms. These 3000 firms should not be compared to the sample size of about 41 000 firms. Rather, they must be compared to the extended groups of firm which would have included all the other firms I have excluded because of poor information. In sum, it must be recognized that the selection criterions above sets a lower limit to the quality of firms included. In order to be included in the analysis, a firm must fulfil some minimum requirements, most notably inclusion in all the relevant registers and reporting accounts from the start, which in and of itself are good signals.

4.2 Variables

The analysis takes a snapshot of the firm in its year of entry, and uses these "first year characteristics" to explain later survival and exit risks. By only using information from the firm's first year I ensure that the explanatory variables used are exogenous or at least predetermined to the exit decision. The only time-varying explanatory variable will therefore be the firm age and calendar year dummy variables. Other variables characterize the firm, its industry and its employees in the year of entry.

4.2.1 Firm characteristics

I want to control for the heterogeneity of firm which is caused by industry, geographical location and listing on the stock exchange. The cross-section data set on the firms is my preferred source of information on these characteristics. If missing here, I check to employment and accounts data. The firms' county is my control for location[3].

For industry controls I use a two-digit NACE code. I further group the industries into sectors according to the same standard of industry classification (Eurostat 1996; SSB 2010a). The sector variable was used instead of industry in a simplified regression.

The organizational form of firms is available from both employment, accounts and firm data. For the purpose of the sample selection described earlier, all I need to know is whether or not a firm is organized as a Ltd. firm, that is, either a private or a public limited company. There are few discrepancies between the organizational type reported in the different sources. In the regressions, I also use a dummy distinguishing between private and public limited firms.

4.2.2 Employment and characteristics of the employees

[3] Norwegian county: *Fylke*

The employment level in the firm's first year is our main indicator for firm size, and we use the characteristics of the employees to say something about the type of labour input used and the technology of the firm. The labour input will also be used in the computations of labour productivity.

When assessing the labour input of each firm, I begin with the start- and end- date of each individual employment-relationship to compute the duration of the employment as a fraction of the calendar year. This is the gross employment per individual employee (G_{ij}). Second, gross employment is adjusted by the reported expected hours worked per week (a_{ij}), to get the contribution to net employment from the individual (E_{ij}). Summing all individuals in firm i, I get a measure of the year's total employment (E_i).

$$E_i = \sum_{j=1}^{n} E_{ij} = \sum_{j=1}^{n} a_{ij} G_{ij} = \sum_{j=1}^{n} \left[\frac{a_{ij}\left(ED_{ij} - SD_{ij}\right)}{365} \right]$$

where j is the individual, i is the firm, ED is end date of employment and SD is start date of employment[4]. Different values have been tried for the fraction of a full time equivalent worked by part-time workers (a_{ij}), I use the fractions (1, 2/3, 1/3) for full time, reduced time and part-time, respectively. When I aggregate employment over all firms in the employment register, I get approximately 1.69 million man-labour years for year 2000. This amounts to 93 percent of the 1.83 million man-labour years reported by Statistics Norway (SSB, 2010c).

I want to construct a measure for yearly labour input (E_i^*) based on the above employment measure. The firm might not be in operation for the whole duration of the calendar year, particularly not in the first and last year of operation. I want to be able to compare the size of different firms in their first year by the scale of labour input, even if one firm starts in January and another in August. For this purpose I divide the above employment measure by the fraction (F_i) of the year in which the firm had employees.

$$F_i = \left[\max(ED_{ij}) - \min(SD_{ij}) \right] / 365$$

4 There are a lot of missing values for start- and end-dates, particularly in the early years of the period. The missing values normally means that the employment started before the current year or continues after the end of the year. Therefore, when a start-date is missing, I set it to January 1st, and when an end-date is missing it is set to December 31st.

For example, if the first employee started at July 1st, I will divide by one half, so the labour measure is doubled. I then get the yearly equivalent - how much employment there would have been if the same employment level was expanded to last for a whole year, E_i^*.

$$E^*_i = (E_i / F_i)$$

This is the employment measure used in the analysis. I split the size of employment in five dummies. Most of the firms in the sample are small in terms of employment in the first year, the median firm has 2.7 man-labour years. Only 17% of the firms have more than five man-labour years of employment according to this definition, and they pull up the average to 5.4 man-labour years.

In addition to the use of employment per se as an explanatory variable, I need a labour input measure for the calculation of labour productivity. For this purpose, it would be counterintuitive to use yearly equivalents. If employment lasted for a few months only, then the accounting values are also corresponding to the activity in a period of less than one year. But I want to make another adjustment from the net firm employment E_i. Many Ltd. firms has some labour input from non-employees. There are often owners who get income from self-employment rather than labour-wages as compensation for their efforts. I will assume one man-labour year extra labour input for each firm from non-employees. Through this step, I avoid the very low employment levels otherwise reported for many firms, which in the labour productivity measure would lead to apparent very high productivity if left uncorrected. L_i is the labour input measure used in the later computation of labour productivity. Thus $L_i = E_i + 1$.

When aggregating demographic characteristics of the employees on the firm level, namely the age, education and sex, I weight the contribution from each employee by their respective fractions of the firm's total employment. The variables describing the mean age and education of employees and the share of female employees are grouped in dummies, and each have a separate dummy for missing values. I also include a separate dummy for the reported sex of the firm's CEO. There are many missing values for this variable, but as I have found that the presence of missing information has some explanatory power I choose to include it anyhow.

4.2.3 Labour productivity and debt

The labour productivity of the firm in its first year is a proxy for initial quality and can tell us something about the prospects for future profitability. The productivity measure used here is based on the gross value generated per unit of labour in the firm. I compute it as profit plus wages, divided by labour input. The valuation and cost of capital and the resulting definition of profits used here will be discussed after a passage on the implications of the financial year.

The financial year

Revenue, expenses and wages are all monetary values obtained from the accounts data. Labour input is computed on the basis of yearly employment data. Before combining the two sources in a productivity measure, it is important to bear in mind that the financial year can deviate from the calendar year in two ways. Its duration is not always twelve months, and it may not start and end at the turn of the year. On the one hand, the duration of a financial year can be up to eighteen months. If a firm starts operating after July 1st, it is customary to include the transactions for the first half year in the accounts for the following calendar year. If it starts June 30th or earlier, it must hand in a separate accounts for the first calendar year according to the eighteen months limit. On the other hand, for some firms it might be more appropriate to separate the years at a different time than at new year. For example, a ski resort will prefer to have the whole winter season united in one accounts, and use the slow summer months rather than early spring to prepare the annual accounts. The firm can then use for example May 1st to April 30th as its standard financial year.

These deviations between financial and calendar years make it more troublesome to link the accounts data to the employment data. In order to reduce this problem, I split each entry according to calendar years and allocate the respective fractions of each entry into the right calendar year. All the accounts have variables for the start- and end-dates of its financial year, except for the years 1995 to 1997 when the start-dates are missing. As a point of departure, the start-dates for these years is set to January 1st. The dates are then adjusted to some extent by general rules based on the end-date of each financial year and the start- and end-dates of employment. Next, I adjust the cases where the same six months has been included in the accounts of two successive financial years, and the cases where there are months that appear to be missing in the accounts. Finally, I use these corrected dates to divide and allocate the entry values of all accounts as described above. The end result is that whenever there are discrepancies between financial and calendar year, the firm's accounts for one calendar year will be a weighted average of the accounts for two or three successive financial years.

The valuation of capital

From the accounts data, I get nominal values of sales revenues, wage costs, material costs and other costs as well as provisions for the write-down and depreciation of capital[5]. I want to compute a productivity measure, preferably a labour productivity measure calculated as the operating income less the cost of capital and other non-labour input, divided by labour input. The cost and size of the capital input is by far the most difficult one to assess on the basis of accounts information. The valuation of a plant's capital stock in the annual accounts is not as straightforward as one might hope for. The capital values listed in the firms' accounts are blurred by a substantial degree of subjectivity in the process of writing off and depreciating capital.

According to the Norwegian Accounting Act (Regnskapsloven, 1998), the book-keeping value of capital is computed as the purchase expenses at the date of acquisition adjusted by some depreciation. According to the International Financial Reporting Standard (IFRS), book-keeping value of capital should equal the expected present market value. Public limited companies in Norway must follow the principles in IFRS, and also those who are planning to go public or for other reasons find it more convenient or appropriate to follow IFRS may choose to do so. Those are the exceptions, the vast majority of private limited firms follow the Norwegian principles. Under the Norwegian system, old capital stocks are systematically undervalued compared to new capital stocks. On average this implies an undervaluation of the capital of old firms compared to new firms, if we assume that the age of capital is related to the age of the firm.

This bias is also present in the cost accounts, where the provisions for writing off and depreciation of capital are included to represent the annual cost of capital input. The choice of how to write-off capital is to a large extent a subjective choice of the firm. Another factor worthy of mentioning is that some firms choose to rent their capital equipment rather than buy and own it. If so, their cost of capital is located under the provision *other costs* in the accounts, as opposed to under the provisions for writing off and depreciation where it would be located had the firm owned its own capital.

Because of the imprecise nature of the different capital measures, I tested two different measures of productivity in the regression. One included capital costs and one excluded it, the

---

[5] The correct accounting expression would be the write-down and depreciation of tangible and intangible fixed assets. I use the term capital, as it is more commonly used in the field of economics.

latter case disregarding both the writing off and depreciation of capital and other costs. The first one, with capital cost, was found to give the model larger explanatory power. Because I only use productivity in the first operating year as an explanatory variable in the analysis, the bias from the age of capital is not likely to be important here. Neither is there an apparent reason to believe that the subjectivity and thus differences in the entries for capital costs should be systematically biased in any particular direction. Besides, leaving out the capital cost altogether deprives the analysis of some information. I have therefore used the full productivity measure with capital costs in the estimations reported in this paper. The operating result equals sales income and other income less wages, material costs, other costs and the depreciation and appreciation of capital. Productivity is defined here as operating result plus wages, divided by employment plus one. All nominal values in the accounts are transformed to real values using CPI (SSB, 2010b) before the productivity is computed. I group the labour productivity variable into ten dummies, one for each decile of the distribution.

Debt

There is a close relationship between bankruptcy and debt. Unsettled claims from creditors is the triggering factor for a bankruptcy petition. The ratio of debt to equity is available directly from the accounts data, and is an indicator of the financial solvency of the firm. A high debt-to-equity ratio means that the firm is highly leveraged, so the debt burden is large. A negative value of the debt-to-equity ratio can only occur if the value of equity is negative. Negative equity is regarded as a sign of very high risk of bankruptcy. Because it is a ratio, it can not be split according to calendar years in the same way as the monetary values. Instead, I use the ratio from the firms' first reported accounts. The debt ratios are grouped in nine dummies, and the two first dummies represent negative values of debt-to-equity.

4.3 Estimation procedure

The final data set contains one observation per firm year. Each observation includes the firm identification number, firm age, calendar year, dummies identifying exit and all the above mentioned variables describing the first year of the firm.

The first three models represent three different ways to handle the unobserved differences between firms. Model 1 disregards the unobserved differences. Model 2 imposes a gamma-distributed random covariate to represent the firm-specific unobserved quality parameter v_i. Model 3 uses the method with non-parametric maximum likelihood estimators, which approximates a discrete distribution of v_i.

Model 4 to 6 are extensions of model 3. Model 4 is a single-risk model like model 3, but with a larger set of covariates. Model 5 and 6 are models with two competing risk. Model 5 separates between the risk of "full exit" and the risk of "half exit", while model 6 separates between bankruptcies and liquidation.

In order to estimate model 2, I use a Stata-program developed by Stephen P. Jenkins and presented in Jenkins (1997). This program estimates a clog-log model with and without added Gamma-distributed unobserved heterogeneity.

When estimating the NPMLE-models, models 3 to 6, I use a programme developed at the Frisch Centre. It is written in R programming language, and was designed specifically for the purpose of finding a manageable and computable means to handle the unobserved heterogeneity with a discrete approximation. This is the method which is thoroughly explained and tested in Gaure et al. (2007), and used for example in Røed and Westlie (2007).

The reference model disregarding the unobserved heterogeneity altogether, model 1, has been estimated both in R and in Stata. In R, this corresponds to the case where the discrete distribution of unobserved heterogeneity has only one support point. In Stata, this is the reported basic model before the assumption of Gamma-distributed unobserved heterogeneity is added. The two methods should be identical, and they do indeed produce identical results.

Stata works more slowly than R, and runs into difficulties when the number of parameters gets very large. In the end I landed on a slightly simplified model for the estimation of model 2 in Stata. For comparability, models 1 and 3 include the same set of covariates. This simplified set of covariates include sector to control for industry, and have no indicator for geographical location. The full set, used in models 4 to 6, use the two-digit NACE codes for industry controls, and include the dummies for county.

When estimating the NPMLE-models, I get estimates for each iteration of the model, until the likelihood is no longer improved by including more mass points in the estimated distribution of unobserved firm characteristics v_i. After the maximum likelihood estimation is completed, I choose the iteration with the lowest reported value for AIC. AIC is the Akaike Information

Criterion, a measure of model fit. When modelling probabilities, we get no residuals and hence no R^2 on which to base a measure of model fit. The AIC used is therefore a function of the log-likelihood function, with a penalty for the number of parameters included. AIC improves (gets smaller) as the likelihood function increases and degrades (gets larger) as the model size increases. The best iterations in these models were the ones with four mass points in model 3 to 5, and six mass points in model 6.

5 The estimation results

This chapter will begin with a brief guide to how the estimation results from the duration models can be understood. I start out by looking at the three single risk models with different assumptions on heterogeneity. Next, I turn to models of competing risks. First is a model separating between what I have called full and half exit, where half exit is the label for firms exits where the plant continues under a new firm. The second competing risks model separate between bankruptcies and liquidation, where liquidation includes all non-bankrupt exits. The coefficients estimated for the observed differences between firms will be discussed after the various models and their respective baseline hazards. At last, I give a numerical example which combines selection effects and duration dependence to produce an average exit hazard.

5.1 Interpreting the results

Tables 3, 5 and 6 report the estimation results. The estimated coefficients β and their standard errors are reported. To interpret the estimates, note that $\exp(\beta)$ expresses the proportional change in the continuous time exit hazard of a firm with this value of the covariate, relative to one similar in all other respects but with the reference value for this covariate.

The reference firm is a construct, defined as the firm for which all the reference dummies are equal to one. In our case, it is a firm that started in Oslo in 1996 with 1-2 employees whose mean age was 30-36 years, with a female share below 40%, productivity near the median etc. All the coefficients for the reference dummies are normalized to one. This means that a dummy variable with coefficient such that $\exp(\beta_i) = 1.15$ is associated with approximately 15% larger exit hazard than the reference firm, while one with $\exp(\beta_i) = 0.70$ corresponds to 30% less than the reference.

The asterisks in the columns "sign. β" indicate whether the estimated coefficient for β is significantly different from 0, or, in other words, whether the implied $\exp(\beta)$ is significantly different from 1. If $0 \notin \left[\beta \pm 1.96 \cdot SE(\beta)\right]$, then two asterisks are reported, indicating that the coefficient is significantly different from zero at the 5% level. Estimates that are significant at the 1% level are marked with three asterisks, at the 5% level with two asterisks, and finally, estimates that are significant at the 10% level is marked with one asterisk. The estimated coefficients for sector, NACE and county are found in the appendix to this paper.

The estimated intercepts in the different models will not be a focus of this discussion. They represent the expected hazard for the reference firm in the third year, and the reference firm is not an easily interpreted concept. There are no observations of firms that satisfy all the reference dummies. Nor is the reference firm an average firm. Furthermore, the discrete approximation of the unobserved quality parameters of the sample firms is very crude, and the point estimates are not reliable. The discussion will instead focus on the direction of proportional increases or decreases in the hazard rate for firms with different covariates than the reference firm.

5.2 Different assumptions on heterogeneity

This section will start with a closer look at the observed exit hazard of firms in the sample. I proceed to a model which controls for the observed heterogeneity of firms (model 1). Finally, I present two models which attempt to filter out the selection effect from the duration dependence by explicitly modelling the unobserved heterogeneity. One model assumes a Gamma-distribution of the unobserved heterogeneity (model 2), the other searches for a discrete distribution (model 3).

The observed gross exit rate in the sample is shown in Figure 1. After a peak of nearly 8% firm exits per year in the third year, the exit hazard declines to about 4% in year nine. As explained in section 4.1.7, the sample probably has a surprisingly low exit rate for firms aged one to two years, partly because only firms that report accounts for at least one year are included in the sample. I will therefore not emphasize the hazard rates for year one and two in the discussion of estimation results, and instead use the third year as the reference year. Nevertheless, it remains an empirical fact that the gross exit hazard for the firms in this sample is increasing until year three. There may be other explanations than sample selection.

		Model 1			Model 2			Model 3		
Log-likelihood				-42428.0257			-42262.476			-42246.6094
Parameters				74			75			80
Mass points				1			-			4
			Basic			**Gamma**			**Discrete**	
Variable	dummy	β	std. β	sign. β	β	std. β	sign. β	β	std. β	sign. β
Firm age	age 1	-0.22800	0.03112	***	-0.71351	0.04727	***	-0.81233	0.06009	***
	age 2	-0.02647	0.02980		-0.25707	0.03432	***	-0.30321	0.03740	***
	age 3 (R)									
	age 4	-0.04536	0.03301		0.15618	0.03644	***	0.19544	0.03838	***
	age 5	-0.05615	0.03621		0.33094	0.04645	***	0.40252	0.05003	***
	age 6	-0.07665	0.04125	*	0.48228	0.05903	***	0.57632	0.06335	***
	age 7	-0.09688	0.04800	**	0.63465	0.07339	***	0.73876	0.07702	***
	age 8	-0.11703	0.06394	*	0.79099	0.09484	***	0.89150	0.09556	***
	age 9	-0.17791	0.10843		0.91526	0.13931	***	0.99973	0.13671	***
Entry type	**new plant (R)**									
	old plant	-0.19162	0.04569	***	-0.31555	0.06456	***	-0.33250	0.06826	***
	old pl, non-Ltd.	-0.47970	0.04150	***	-0.73606	0.05836	***	-0.78060	0.06257	***
	old pl, multi-pl	-0.29459	0.10172	***	-0.46482	0.13515	***	-0.49871	0.14428	***
	old pl, non-Ltd. Mult	-0.15278	0.19492		-0.17465	0.27212		-0.18348	0.28969	
Year	1996	-0.24507	0.06991	***	-0.23987	0.07458	***	-0.23810	0.07569	***
	1997	-0.12556	0.04837	***	-0.13935	0.05035	***	-0.14303	0.05079	***
	1998 (R)									
	1999	0.03908	0.03975		0.06870	0.04120	*	0.07337	0.04150	*
	2000	0.19858	0.03854	***	0.26334	0.04166	***	0.27111	0.04255	***
	2001	0.11782	0.03958	***	0.20526	0.04416	***	0.21474	0.04544	***
	2002	0.19103	0.03931	***	0.29605	0.04570	***	0.30763	0.04751	***
	2003	0.01555	0.04095		0.11896	0.04877	**	0.13025	0.05085	**
	2004	-0.47809	0.04574	***	-0.42510	0.05391	***	-0.41922	0.05676	***
Labour productivity	1st decile	0.31804	0.03943	***	0.49092	0.06254	***	0.53144	0.06477	***
	2nd decile	0.32745	0.03807	***	0.47720	0.05935	***	0.49296	0.06327	***
	3rd decile	0.14881	0.03852	***	0.21998	0.05865	***	0.23153	0.06172	***
	4th decile	0.10378	0.03876	***	0.12705	0.05863	**	0.14621	0.06153	**
	5th decile (R)									
	6th decile	-0.08310	0.04105	**	-0.13916	0.06023	**	-0.14367	0.06280	**
	7th decile	-0.19276	0.04313	***	-0.30693	0.06241	***	-0.31994	0.06553	***
	8th decile	-0.29806	0.04613	***	-0.42259	0.06527	***	-0.43501	0.06787	***
	9th decile	-0.21782	0.04612	***	-0.32731	0.06541	***	-0.34193	0.06897	***
	10th decile	-0.20386	0.04709	***	-0.31764	0.06698	***	-0.33814	0.07135	***
Debt-to-equity	< -, -5]	0.38585	0.03534	***	0.76277	0.05948	***	0.81581	0.05985	***
	<-5, 0]	0.62840	0.03695	***	1.14657	0.06569	***	1.21544	0.06811	***
	<0, 0.5]	0.00893	0.05276		0.01979	0.07574		0.02420	0.08202	
	<0.5, 1]	-0.00536	0.05002		0.01371	0.07034		0.01478	0.07588	
	<1, 2.5] (R)									
	<2.5, 5]	-0.03085	0.03418		-0.02064	0.04793		-0.01696	0.05170	
	<5, 10]	-0.06878	0.03478	**	-0.06410	0.04859		-0.06404	0.05225	
	<10, 100]	0.15815	0.03490	***	0.27790	0.05147	***	0.30572	0.05496	***
	<100, +]	0.24138	0.08479	***	0.40719	0.12947	***	0.44930	0.13703	***
	missing	0.37775	0.12486	***	0.81549	0.20966	***	0.85379	0.20584	***
Employment	<0, 1]	-0.04291	0.02584	*	-0.05495	0.03901		-0.06289	0.04165	
	<1, 2] (R)									
	<2, 3]	-0.03202	0.02964		-0.01123	0.04460		-0.02008	0.04655	
	<3, 5]	-0.02185	0.03059		0.00580	0.04605		0.00562	0.04754	
	<5, +]	-0.00193	0.03225		0.02436	0.04784		0.01783	0.04928	
Age of employees	<-, 30>	0.11287	0.02456	***	0.18978	0.03826	***	0.19547	0.03975	***
	[30, 36> (R)									
	[36, 42>	-0.09302	0.02628	***	-0.16936	0.03925	***	-0.18345	0.04136	***
	[42, +>	-0.19315	0.02705	***	-0.31935	0.04064	***	-0.33502	0.04287	***
	missing	0.26258	0.21471		0.63938	0.34267	*	0.61699	0.32870	*
Education of employees	<-, 10>	0.07167	0.03606	**	0.14747	0.05799	**	0.13776	0.05989	**
	[10, 12> (R)									
	[12, 14>	-0.09791	0.02203	***	-0.16908	0.03349	***	-0.17672	0.03511	***
	[14, +]	-0.24542	0.03275	***	-0.35978	0.04794	***	-0.38381	0.05110	***
	missing	-0.02737	0.06595		-0.03582	0.10228		-0.03759	0.10619	
Female share of employees	**[0, 0.4] (R)**									
	<0.4, 0.6>	-0.02547	0.03261		-0.04429	0.04953		-0.04848	0.05152	
	[0.6, 1]	-0.02249	0.02553		-0.06189	0.03937		-0.06053	0.04145	
	missing	0.05043	0.15283		0.13520	0.24038		0.16160	0.24871	
Sex of CEO	**male (R)**									
	female	-0.00057	0.03215		0.03983	0.04830		0.04982	0.04987	
	missing	0.34158	0.02154	***	0.48960	0.03464	***	0.52161	0.03840	***
Public vs private Ltd.	**private (R)**									
	public	-0.13110	0.17534		-0.17090	0.26223		-0.19355	0.26676	

Table 3 - Estimation results for models 1, 2 and 3

For example, it might be that many entrepreneurs decide to “dig in” and keep the business running for at least a couple of years before they cave in, even if they are losing money. Having gone through the efforts of establishing a new firm, they might want to "give it a real try".

Controlling for unobserved heterogeneity in the regression is an attempt to separate the effect of duration from the effect of type. I compare three different model specifications. Model 1 disregards the unobserved heterogeneity and estimates the gross effect of age. This model, therefore, combines selection effects and duration effects. If there were no unobserved heterogeneity there would be no remaining selection effect once the observed characteristics were controlled for, and model 1 would be sufficient. Model 2 assumes that the unobserved heterogeneity is Gamma-distributed. The mean of the distribution is set to one, and the regression gives an estimate of the variance. This estimate provides an indication of the role played by unobserved heterogeneity in the data. Model 3 estimates a discrete distribution of the unobserved heterogeneity as explained in chapter 3.2. The estimation results of these three regressions are shown in Table 3.

As we can see in Figure 2b, model 1 has no significant effects of age after year three. If anything the tendency is towards a slightly decreasing hazard. Compared to the observed gross exit rate shown in Figure 1, this basic model removes the part of the selection effect that stems from observed differences between firms. The remainder of the selection effect, the one caused by quality difference we have not been able to observe and include as covariates in the model, remains.

After controlling for observed heterogeneity in model 1, the selection effect was clearly reduced compared to the effect in the observed exit hazard. Though the selection effect must still be assumed to be significant, it is no longer strong enough to create a downward sloping hazard function.

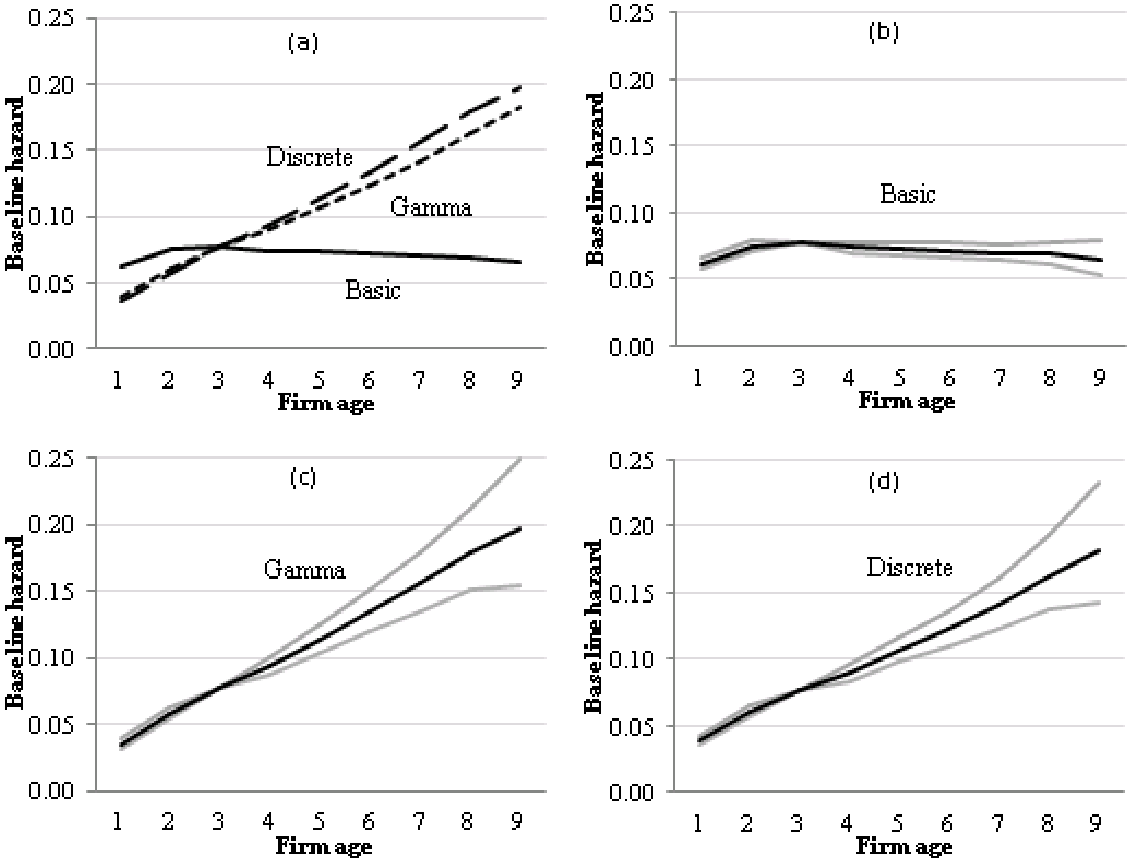

Figure 2 - Baseline hazards estimated in models 1, 2 and 3
(a) Models 1, 2 and 3 combined (b) Model 1, basic
(c) Model 2, Gamma (d) Model 3, Discrete

Model 2 gives an estimated duration dependency increasing in age. The estimated variance of the unobserved heterogeneity ν is significant at a 1 percent level. That indicates significant unobserved heterogeneity in the sample, which would imply that a selection effect is included in the estimated duration dependence of model 1. Model 2 and model 3 give very similar estimates, in fact there are no substantial differences between the estimated coefficients β in the two models. This indicates that the true distribution of the unobserved heterogeneity in the sample might not be very different from the Gamma-distribution.

Figure 2 illustrates the duration dependence with different assumptions regarding the unobserved heterogeneity of firms. The observed exit frequency of all exits at age three is used to give a reference level, and the other years are scaled according to the estimated coefficients λ_t in the different models. If observed exit in year three is h_3, then ν is given by

$h_3 = 1 - \exp(-\exp(\hat{\nu}))$, and h_t is given by $h_t = 1 - \exp(-\exp(\lambda_t + \hat{\nu}))$. Figure 2b-d include a 95% confidence interval for the estimated duration dependence.

From Table 3, we see that the calendar year affect is remarkably unaffected by the inclusion of ν in models 2 and 3 compared to model 1. The remainder of the explanatory variables in the models differ only in terms of scale between the different models. Note that the employment level in the year of entry is not significant in these models. Firms that start up with older, more educated employees have a lower exit hazard, and so do firms that start up with an old plant. A further discussion of the significance of observed differences between firms will follow in chapter 5.4.

5.3 Competing risks

All the firm exits in the sample are categorized as either full or half exit, and as either bankruptcy or liquidation, in total four different possible combinations. For example, 4% of the exits are bankrupt firms whose plant continues under a new firm. Table 4 reports the total frequencies of different types of exit. The first competing risks model (model 5) separates the 14% "half exits", in which the plant continues after firm exit, form the "full exits". The second competing risks model (model 6) separates the 49% bankruptcies from the remaining firm exits.

Type of exit	Bankruptcy	Liquidation	Total
Full exit	5519	4984	10503
	45%	41%	**86%**
Half exit	536	1196	1732
	4%	10%	**14%**
Total	6055	6180	12235
	49%	**51%**	100%

Table 4 - Four different types of exit

The motivation for splitting the exits in different groups is that the observed duration dependencies are quite different for the different types of exit. As we have seen, the observed exit rate is decreasing with firm age after year three. Figure 3a shows a decomposition of the gross exit rate into full and half exits. The hazard of half exit, that is, plant continuation, is increasing with age even before I correct for the selection effect. Figure 3b shows a similar decomposition, now separating bankruptcies from liquidation. The observed risk of bankruptcy is steadily falling with firm age, but there is no clear trend in the risk of liquidation after the third year.

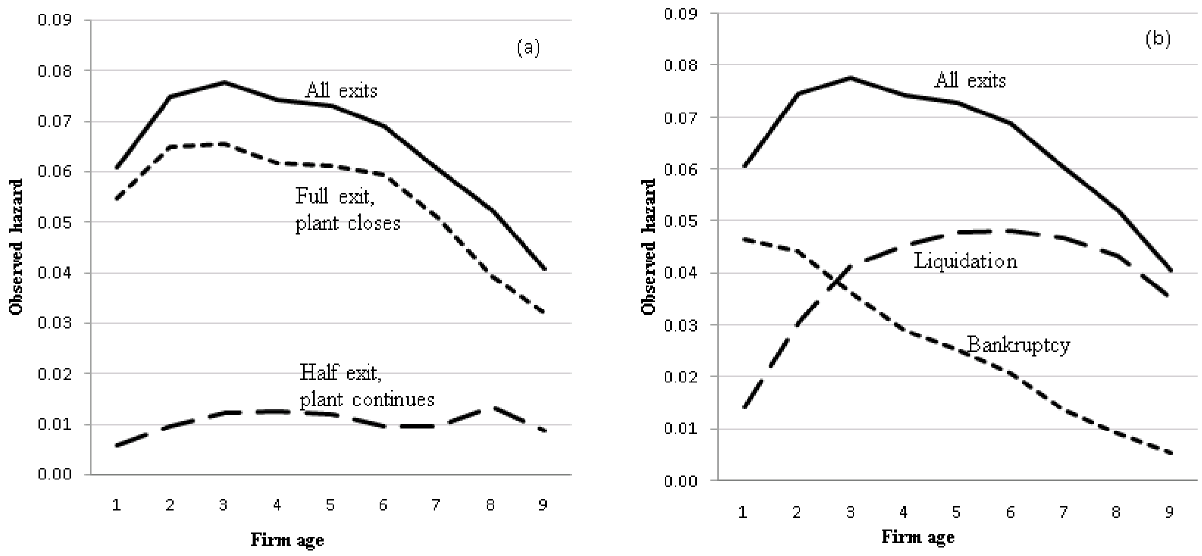

Figure 3 - Observed exit decomposed

The estimated duration dependencies in the two competing risks models can be compared to a single risk model. To do this, a model 4 was estimated, identical to model 3 but with sector controls replaced by the more detailed information in NACE-category and county. Model 4 was also estimated using two different samples. First, a sample excluding firms with "half entry", where the plant is not new. Second, a sample excluding firms observed with employees the year before their first reported accounts. The duration dependence estimated from these samples turned out in essence identical to the results in the full sample.

I will first present the estimated duration dependence of the different types of exit, and then turn to the varying impact of the other explanatory variables in the models.

		Model 4			Model 5					
Log-likelihood				-42063.1363						-46047.1839
Parameters				139						265
Mass points				4						4
		All exits			Full exit			Half exit		
Variable	dummy	β	SE(β)	sign. β	β	SE(β)	sign. β	β	SE(β)	sign. β
Firm age	age 1	-0.81901	0.05901	***	-0.79866	0.06080	***	-1.35044	0.13682	***
	age 2	-0.30242	0.03710	***	-0.29301	0.03924	***	-0.51695	0.09567	***
	age 3 (R)									
	age 4	0.19017	0.03811	***	0.19804	0.04120	***	0.26824	0.09385	***
	age 5	0.38973	0.04931	***	0.41641	0.05257	***	0.46668	0.11818	***
	age 6	0.55597	0.06217	***	0.61728	0.06608	***	0.51689	0.14610	***
	age 7	0.71020	0.07495	***	0.76276	0.08008	***	0.79361	0.17090	***
	age 8	0.85502	0.09327	***	0.80628	0.10194	***	1.35331	0.19815	***
	age 9	0.95815	0.13478	***	0.98699	0.14847	***	1.23841	0.29254	***
Entry type	**new plant (R)**									
	old plant	-0.34226	0.06840	***	-0.56062	0.07725	***	0.17206	0.10132	*
	old pl, non-Ltd.	-0.77448	0.06286	***	-0.91304	0.06866	***	-0.29478	0.11031	***
	old pl, multi-pl	-0.54911	0.14506	***	-0.90647	0.17884	***	0.03685	0.19731	
	old pl, non-Ltd. Multipl	-0.30135	0.29482		-0.56847	0.33917	*	0.07265	0.43099	
Year	1996	-0.27211	0.07712	***	-0.26219	0.07948	***	-0.77653	0.34424	**
	1997	-0.15092	0.05106	***	-0.14376	0.05326	***	-0.35175	0.18791	*
	1998 (R)									
	1999	0.07886	0.04169	*	0.04926	0.04436		0.37019	0.12493	***
	2000	0.28046	0.04282	***	0.25322	0.04548	***	0.58348	0.12404	***
	2001	0.22762	0.04570	***	0.19172	0.04855	***	0.58354	0.12645	***
	2002	0.32369	0.04776	***	0.31104	0.05053	***	0.56115	0.13086	***
	2003	0.14769	0.05111	***	0.13122	0.05427	**	0.42069	0.13406	***
	2004	-0.39917	0.05693	***	-0.47126	0.06127	***	0.08107	0.14136	
Labour productivity	1st decile	0.53813	0.06474	***	0.56273	0.06849	***	0.56824	0.12841	***
	2nd decile	0.49110	0.06297	***	0.55324	0.06669	***	0.24894	0.12451	**
	3rd decile	0.21992	0.06141	***	0.26903	0.06532	***	-0.05752	0.12010	
	4th decile	0.14419	0.06114	**	0.13582	0.06494	**	0.23314	0.11453	**
	5th decile (R)									
	6th decile	-0.13804	0.06248	**	-0.13208	0.06625	**	-0.18528	0.12366	
	7th decile	-0.31175	0.06535	***	-0.32312	0.06959	***	-0.30052	0.12625	**
	8th decile	-0.42569	0.06778	***	-0.43113	0.07244	***	-0.36708	0.13474	***
	9th decile	-0.32570	0.06920	***	-0.34498	0.07367	***	-0.15801	0.14102	
	10th decile	-0.31379	0.07186	***	-0.35818	0.07708	***	0.11006	0.14286	
Debt-to-equity	< -, -5]	0.81810	0.05950	***	0.83553	0.06277	***	1.05764	0.13292	***
	<-5, 0]	1.20116	0.06733	***	1.25874	0.07114	***	1.20844	0.15075	***
	<0, 0.5]	0.05622	0.08224		0.01831	0.08701		0.46961	0.20457	**
	<0.5, 1]	0.01609	0.07603		-0.00416	0.08011		0.21472	0.19723	
	<1, 2.5] (R)									
	<2.5, 5]	-0.01896	0.05180		-0.02395	0.05457		0.12646	0.12380	
	<5, 10]	-0.05842	0.05233		-0.10913	0.05565	**	0.33755	0.11773	***
	<10, 100]	0.31733	0.05494	***	0.28159	0.05819	***	0.70153	0.11965	***
	<100, +]	0.48186	0.13711	***	0.42397	0.14875	***	0.96387	0.23263	***
	missing	0.83025	0.20891	***	0.85071	0.21942	***	1.25893	0.38293	***
Employment	<0, 1]	-0.02912	0.04169		-0.02088	0.04373		-0.41210	0.10808	***
	<1, 2] (R)									
	<2, 3]	-0.03444	0.04655		-0.07144	0.04909		0.28599	0.09762	***
	<3, 5]	-0.01493	0.04763		-0.09938	0.05096	*	0.53317	0.09380	***
	<5, +]	0.01120	0.04941		-0.23000	0.05412	***	1.00967	0.09420	***
Age of employees	<-, 30>	0.17167	0.03958	***	0.16849	0.04201	***	0.20121	0.07598	***
	[30, 36> (R)									
	[36, 42>	-0.17427	0.04143	***	-0.19429	0.04423	***	-0.10571	0.08328	
	[42, +>	-0.30873	0.04298	***	-0.30680	0.04550	***	-0.44657	0.09631	***
	missing	0.75192	0.32626	**	0.82239	0.34413	**	-0.62526	1.18921	
Education of employees	<-, 10>	0.10996	0.05999	*	0.11290	0.06373	*	0.14000	0.11868	
	[10, 12> (R)									
	[12, 14>	-0.18293	0.03537	***	-0.21218	0.03772	***	-0.03105	0.07025	
	[14, +]	-0.44410	0.05223	***	-0.47867	0.05542	***	-0.23648	0.12409	*
	missing	-0.09722	0.10673		-0.10205	0.11182		-0.08770	0.29859	
Female share of employees	**[0, 0.4] (R)**									
	<0.4, 0.6>	-0.08489	0.05211		-0.15226	0.05610	***	0.19600	0.09184	**
	[0.6, 1]	-0.13596	0.04319	***	-0.23454	0.04638	***	0.28815	0.07982	***
	missing	0.12847	0.24689		0.11682	0.25991		0.59811	0.56732	
Sex of CEO	**male (R)**									
	female	0.01298	0.04996		0.05904	0.05379		-0.08189	0.08863	
	missing	0.51237	0.03808	***	0.54040	0.04030	***	0.41934	0.07675	***
Public vs private Ltd.	**private (R)**									
	public	-0.32130	0.27599		-0.29652	0.29102		0.01608	0.80543	

Table 5 - Estimation results for models 4 and 5

5.3.1 Full exit vs. half exit

Not all exits are failures. Some firms exit even thought they are profitable. What I have labelled a "half exit" is defined as a firm exit in which the plant continues afterwards under another firm. This could happen as part of a change in organizational form, a merger or a firm acquisition. Even when the firm is bankrupt, the plant could be sufficiently profitable that some other firm buys the bankrupt estate in order to continue operations.

The estimation results for model 5 in Table 5 indicates that the positive duration dependence of half exits is more pronounced than that of full exits. This is consistent with the observed hazards shown in Figure 3a, where the risk of half exit is initially very low, gradually increasing with firm age even before the selection effect is accounted for. Figure 4 is constructed in the same way as Figure 2 - Baseline hazards estimated in models 1, 2 and 3Figure 2, with the observed exit rates for each type of exit in year three as the scaling factor.

5.3.2 Bankruptcy vs. liquidation

Bankruptcies are quite different from voluntary firm liquidations. As can be seen from the estimation results for model 6 in Table 6, there are significant differences between the exits due to bankruptcy and the voluntary liquidation of the firm. Both the duration dependence and the estimated effect of explanatory variables differ notably.

The risk of bankruptcy is not increasing with firm age. The estimates suggest a slight decline in the bankruptcy hazard after the third year. The risk of liquidation, on the other hand, is significantly increasing with firm age. This implies that the upward slope estimated for the aggregate duration dependence in model 4 is driven by liquidations. Further evidence of this can be seen from Figure 3b, which shows us that liquidations constitute an increasing share of total firm exits as time goes by.

		Model 4			Model 6					
Log-likelihood				-42063.1363						-48442.9693
Parameters				139						274
Mass points				4						6
		All exits			**Bankruptcy**			**Liquidation**		
Variable	dummy	β	SE(β)	sign. β	β	SE(β)	sign. β	β	SE(β)	sign. β
Firm age	age 1	-0.81901	0.05901	***	-0.49066	0.11410	***	-1.22344	0.06307	***
	age 2	-0.30242	0.03710	***	-0.05211	0.05464		-0.40743	0.04746	***
	age 3 (R)									
	age 4	0.19017	0.03811	***	-0.08372	0.05589		0.22754	0.04754	***
	age 5	0.38973	0.04931	***	-0.13171	0.06973	*	0.44379	0.05743	***
	age 6	0.55597	0.06217	***	-0.17218	0.08670	**	0.60382	0.07005	***
	age 7	0.71020	0.07495	***	-0.25104	0.11487	**	0.73773	0.08257	***
	age 8	0.85502	0.09327	***	-0.23350	0.16667		0.82935	0.10449	***
	age 9	0.95815	0.13478	***	-0.03272	0.30924		0.84219	0.14968	***
Entry type	**new plant (R)**									
	old plant	-0.34226	0.06840	***	-0.44207	0.10246	***	-0.18265	0.08572	**
	old pl, non-Ltd.	-0.77448	0.06286	***	-0.96709	0.10355	***	-0.45381	0.07549	***
	old pl, multi-pl	-0.54911	0.14506	***	-0.82397	0.22823	***	-0.12505	0.17691	
	old pl, non-Ltd. Multipl	-0.30135	0.29482		-0.35296	0.44043		-0.31664	0.38320	
Year	1996	-0.27211	0.07712	***	-0.44997	0.10183	***	-0.15994	0.15134	
	1997	-0.15092	0.05106	***	-0.23018	0.06657	***	-0.10389	0.08992	
	1998 (R)									
	1999	0.07886	0.04169	*	0.08827	0.05619		0.15533	0.06533	**
	2000	0.28046	0.04282	***	0.38470	0.05986	***	0.31325	0.06511	***
	2001	0.22762	0.04570	***	0.44033	0.06627	***	0.20310	0.06781	***
	2002	0.32369	0.04776	***	0.47645	0.07067	***	0.37849	0.06854	***
	2003	0.14769	0.05111	***	-0.00035	0.07724		0.40946	0.07074	***
	2004	-0.39917	0.05693	***	-1.25589	0.09713	***	0.13122	0.07576	*
Labour productivity	1st decile	0.53813	0.06474	***	0.50378	0.09221	***	0.56614	0.08147	***
	2nd decile	0.49110	0.06297	***	0.51538	0.08990	***	0.44352	0.07992	***
	3rd decile	0.21992	0.06141	***	0.32781	0.08713	***	0.10847	0.07894	
	4th decile	0.14419	0.06114	**	0.11803	0.08667		0.14241	0.07875	*
	5th decile (R)									
	6th decile	-0.13804	0.06248	**	-0.26336	0.09153	***	0.01770	0.08004	
	7th decile	-0.31175	0.06535	***	-0.42533	0.09698	***	-0.12640	0.08315	
	8th decile	-0.42569	0.06778	***	-0.78181	0.10925	***	-0.07665	0.08466	
	9th decile	-0.32570	0.06920	***	-0.74552	0.11269	***	0.04340	0.08524	
	10th decile	-0.31379	0.07186	***	-1.09002	0.13199	***	0.18988	0.08652	**
Debt-to-equity	< -, -5]	0.81810	0.05950	***	1.48627	0.10878	***	-0.18250	0.07533	**
	<-5, 0]	1.20116	0.06733	***	1.86422	0.12504	***	0.22438	0.08119	***
	<0, 0.5]	0.05622	0.08224		-0.70415	0.15388	***	0.39986	0.09502	***
	<0.5, 1]	0.01609	0.07603		-0.29144	0.12746	**	0.21065	0.08843	**
	<1, 2.5] (R)									
	<2.5, 5]	-0.01896	0.05180		0.21229	0.08032	***	-0.18479	0.06305	***
	<5, 10]	-0.05842	0.05233		0.21113	0.08124	***	-0.27087	0.06388	***
	<10, 100]	0.31733	0.05494	***	0.66734	0.08831	***	-0.05983	0.06690	
	<100, +]	0.48186	0.13711	***	0.72153	0.20395	***	0.10489	0.17062	
	missing	0.83025	0.20891	***	0.48191	0.32944		0.78452	0.21215	***
Employment	<0, 1]	-0.02912	0.04169		-0.00550	0.06112		-0.01407	0.05125	
	<1, 2] (R)									
	<2, 3]	-0.03444	0.04655		0.03841	0.06671		-0.10063	0.05913	*
	<3, 5]	-0.01493	0.04763		0.02389	0.06841		-0.05680	0.06106	
	<5, +]	0.01120	0.04941		-0.17062	0.07358	**	0.18380	0.06272	***
Age of employees	<-, 30>	0.17167	0.03958	***	0.26246	0.05642	***	0.02390	0.05077	
	[30, 36> (R)									
	[36, 42>	-0.17427	0.04143	***	-0.25584	0.06190	***	-0.07566	0.05162	
	[42, +>	-0.30873	0.04298	***	-0.62437	0.07113	***	-0.02065	0.05221	
	missing	0.75192	0.32626	**	1.29787	0.44554	***	-0.46200	0.46791	
Education of employees	<-, 10>	0.10996	0.05999	*	0.24910	0.08383	***	0.05117	0.07631	
	[10, 12> (R)									
	[12, 14>	-0.18293	0.03537	***	-0.38782	0.05372	***	0.06648	0.04492	
	[14, +]	-0.44410	0.05223	***	-1.11829	0.09852	***	0.01085	0.06130	
	missing	-0.09722	0.10673		-0.24959	0.15346		0.11858	0.13166	
Female share of employees	**[0, 0.4] (R)**									
	<0.4, 0.6>	-0.08489	0.05211		-0.19942	0.07658	***	0.02729	0.06520	
	[0.6, 1]	-0.13596	0.04319	***	-0.37022	0.06490	***	0.08858	0.05429	
	missing	0.12847	0.24689		0.43646	0.33560		-0.32919	0.32086	
Sex of CEO	**male (R)**									
	female	0.01298	0.04996		0.05877	0.07107		-0.03605	0.06373	
	missing	0.51237	0.03808	***	-0.24399	0.05574	***	1.13840	0.05097	***
Public vs private Ltd.	**private (R)**									
	public	-0.32130	0.27599		0.06859	0.40706		-0.32668	0.34288	

Table 6 - Estimation results for models 4 and 6

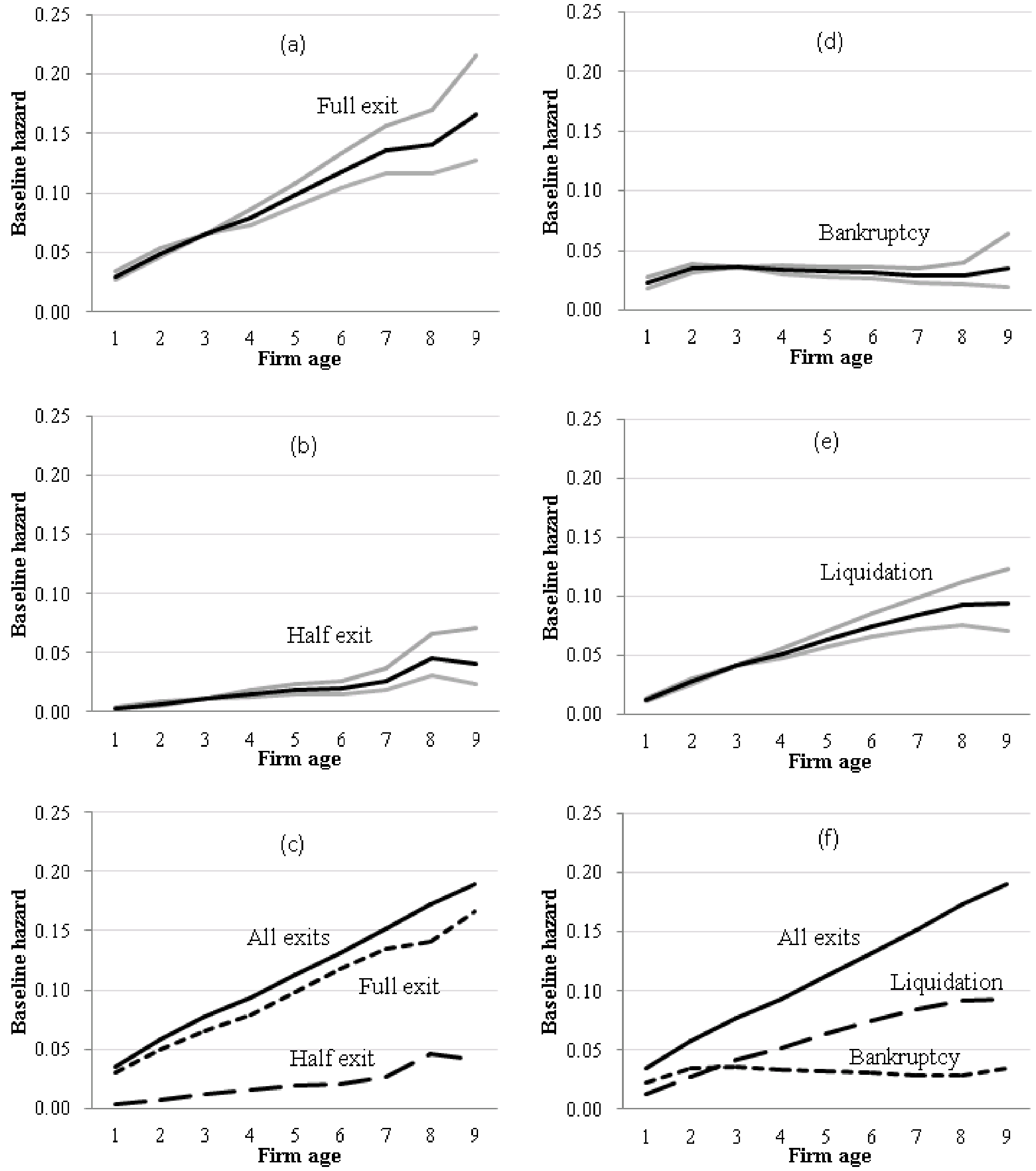

Figure 4 - Baseline hazards estimated in models 4, 5 and 6

(a) Model 5, full exit
(b) Model 5, half exit
(c) Models 4 and 5 combined
(d) Model 6, bankruptcy
(e) Model 6, liquidation
(f) Models 4 and 6 combined.

5.4 The observed heterogeneity of firms

Type of entry

Starting a firm with an already established plant seems to be associated with a lowered exit hazard. At first glance, this may seem puzzling. The capital vintage mechanism, for instance, would imply that the age of the plant increases the exit hazard, not just the age of the firm. However, this ignores the positive selection of surviving plants. An old plant that continues has already survived for some time and proved viable (e.g., as implied by the theory of passive learning). More interesting, however, is the finding that the lower exit rate of firms started with old plants to a large extent reflects a decreased bankruptcy hazard. Bankruptcy does not exhibit a positive duration dependence. A firm with an old plant would therefore seem to have a lower aggregate probability of exit mainly because it has a lower probability of bankruptcy. The significant coefficients for old plants, in addition to the differences discussed between full and half exits, supports the separation of the histories of firm and plant.

Employees

The estimated coefficients for the *size of employment in the first year* are not significant when we look at all exits, such as in models 2 to 4. The reason is that the coefficients on employment have opposite signs for different types of exit, which cancel out if we consider all exits together. From model 5, we see that the firms with high employment have a *reduced* risk of full exit, while they have an *increased* risk of half exit. Similarly, from the estimation of model 6 I find that having more than five employees is associated with a *higher* risk of liquidation, but a *lower* risk of bankruptcy.

A firm with *older and more educated employees* has a significantly lower risk of exit than the reference firm. This holds for all types of exits except liquidation, for which there are no significant coefficients for age and education of employees. These coefficients on age and education must be seen in relation to labour productivity, as the labour productivity measure includes wages. Older and more educated workers generally earn more, which in itself should be reducing the profitability of the firm. On the other hand, the higher wage is also assumed to reflect higher average productivity relative to inexperienced and uneducated workers.

The firms with a *large share of female employees* is found to have a lower risk of exit than firms with more men. As with the other estimated coefficients, we should be careful not to assume a causal effect, in this case of women, on the exit hazard. Still, it is very interesting to see that a large share of female employees in the first year is associated with a significantly lower risk of bankruptcy but no higher risk of liquidation. The risk of bankruptcy is about 30 % lower for a firm with a female share above 60% than for a firm with female share below 40%. If women are more risk averse than men, as discussed in the literature on financial decision making (Jianakoplos and Bernasek, 1998; Schubert et al., 1999; Eckel and Grossman, 2003), this could be part of the explanation. Female employees may then influence the firm to take on less risk, in particular, less debt. Risk averse females may also tend to seek out less risky jobs. Rather than influencing the risk-affecting behaviour of their workplace, they may then serve as an indicator of more stable firm prospects. A result consistent with this is that the risk of half exit, which means a continuation of the plant and thus probably of most of the jobs as well, is found to be higher for firms with a high share of female employees.

Productivity and debt

Negative *debt-to-equity* in the first year is, as expected, strongly linked to a high risk of bankruptcy. The firms in the very worst category of negative equity have a bankruptcy hazard four to five times larger than the reference, and they even have a reduced probability of liquidation. Firms starting up with a debt burden this heavy seem more likely to exit through bankruptcy rather than liquidation. This seems plausible: Unless the debt-to-equity improves before exit, the creditors will demand a bankruptcy process. 20% of the firms in my sample has a negative debt-to-equity in their first year. These firms make up 40% of the bankruptcies. High values of debt-to equity is also associated with a higher total risk of exit, raising the risks of all three types of exit other than liquidation.

High *initial labour productivity* is associated with a significantly lower probability of exit. This is in accordance with both theory and simple intuition: High labour productivity and high profitability indicates a healthy firm. As with many of the other covariates discussed above, the coefficients on productivity are larger and more significant for bankruptcy than for liquidation. This is probably related to the fact that there are very diverse reasons and motivations for liquidation, while bankruptcy is a clearly defined process, initiated because the firm is unable to pay its creditors.

Business cycle and sector shift

As mentioned in the introduction, sector shifts and business cycle conditions also affect exit and entry rates. These are not the focus of this paper, but their effects have been controlled for by year and sector dummies. The results show that there are significant differences between the exit hazards for different sectors and geographical locations (see Tables 9-11 in the appendix). Firms located in the two northernmost counties in Norway have a higher risk of bankruptcy and a correspondingly lower risk of liquidation than do Oslo-based firms.

Differences between sectors may both reflect persistent differences in the average lifetime of firms in the sectors, and sector shifts. For example, hotels and restaurants (sector h) have a significantly higher average exit rate than firms in the retail sector (sector g), while the exit hazard of firms in health and social work (sector n) is lower.

The significantly large and negative coefficient on the risk of bankruptcy in year 2004 is most likely a result of the registration lag explained in section 4.1.7. This demonstrates the value of a calendar control; because it is included in a calendar year dummy, the poor information on bankruptcy in 2004 does not affect the estimated duration dependency.

5.5 Recomposing the exit hazard

The reported mass points in the NPMLE-models, with estimated intercept hazards v_i and assigned probabilities p_i, can not be interpreted as representing real groups of firms, they are rather a construct to assess the approximate size of the selection effect. Nevertheless, I will use the estimates of intercept values and assigned probability for each of the four types of heterogeneity estimated in model 4 in a thought experiment. Combined with the estimated duration dependence from the same model, I will illustrate the selection effect and the duration dependence in a numerical example based on estimated parameters. By combining the effect of type and the effect of age, I can reconstruct the gross exit rate.
The exit hazard for the reference firm was evaluated with all the dummies in the x-vector of explanatory variables equal zero. The hazard function for a reference firm of type i in period t can therefore be simplified to

$$h_{it}(\gamma_t, \nu_i) = 1 - \exp\left\{-\exp\left[\gamma_t + \nu_i\right]\right\}$$

In order to isolate the duration dependence, consider a case where all firms are of the same type. They all have the same exit hazard, starting at the average exit hazard observed in year one and scaled according to the estimated duration dependence from year 2 and onwards. This correspond to setting the same value of ν for all firms, in the example so that the average exit rate is 6.8% for the first year. The development of the hazard rate of these firms is depicted in Figure 5. This is also the shape of the time path of the exit hazard for an individual firm with an initial exit hazard equal to the average. Figure 5Figure 5 illustrates the positive duration dependence of the hazard function for firm exits.

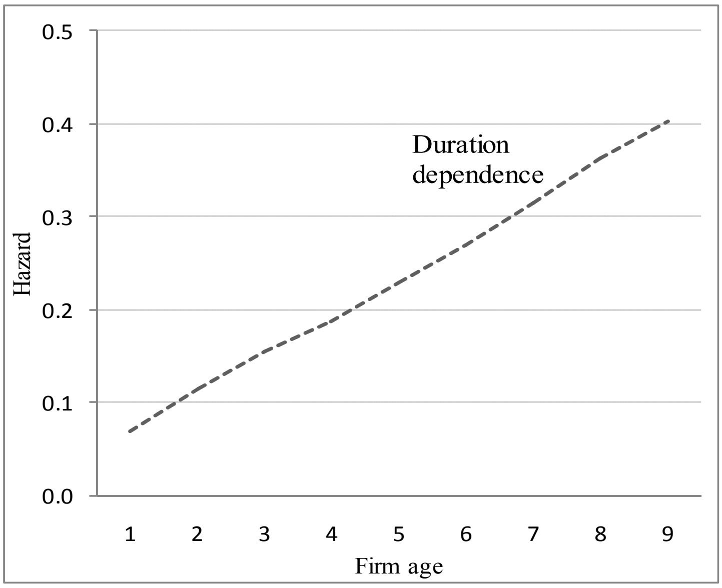

Figure 5 - Duration dependence

Next, to isolate the selection effect, consider a case where the exit rate is constant for any firm, and where the firms can be divided into four groups with different quality parameters. The groups can be described by the intercept hazard ν_i (reflecting "quality") and the probability p_i reported for the four mass points in the distribution estimated for model 4 (reflecting their relative size). Each group has a time-constant exit hazard, given by the reported intercepts, and the share of firms in each group corresponds to the estimated probability of each unobserved heterogeneity mass point. This corresponds to setting the duration dependence λ_t equal to 0 for all time periods. The average hazard rate of these firms would be decreasing in time, because the firms with a low hazard rate makes up an increasing share of the surviving firms. Figure 6 depicts the development of the gross hazard rate, in this case a result of a selection effect operating on unobserved differences.

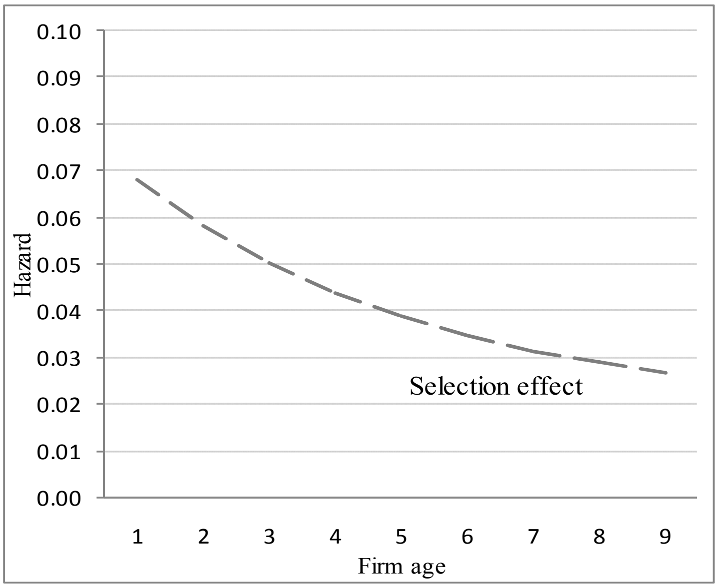

Figure 6 - Selection effect

Finally, consider a more complex world, in which there are four types of firms, with four different quality parameters v_i, each group with the same positive duration dependence given by the vector of λ_t. In this case, the duration dependence and the selection effect operates simultaneously. This corresponds to a weighted average in each time period $t = 1,...,9$ of the four hazard functions

$$h_{ti}(\lambda_t, v_i) = 1 - \exp\left\{-\exp\left[\lambda_t + v_i\right]\right\} \quad \text{where} \quad i = 1,...,4.$$

In period 1 the weights are given by p_i, but from period 2, the weights change because the high quality firms constitute an increasing share of the sample. For each year, the weight assigned to group *i*'s hazard is the share of type *i* firms remaining in the sample. The resulting gross hazard rate is shown in Figure 7. This gross hazard rate closely resembles the hazard curve for model 1 depicted in Figure 2b, in which only observed heterogeneity is accounted for. It is upward sloping until a peak in year three before the average hazard slowly decreases. In the beginning, the effect of duration dependence dominates the gross development and produces an increasing gross exit rate. After the peak at age three, the selection process starts to dominate as more and more of the low-quality firms have left the sample. Even though this exercise should not be interpreted as a realistic description of the population of firms, it illustrates how the opposing forces of type and duration can create gross hazard rates similar to the ones we observe in the sample. It may even be that the observed increasing exit hazard from year 1 to year 3 is not just a result of sample selection: The observed exit hazard also includes the selection effect from *observed* heterogeneity, i.e. differences in firm quality associated with observable characteristics. This creates a stronger downward bias on the effect of age on firm exit compared to what we saw in the above illustration. If duration dependence in the population of firms is a stronger effect than selection in the first couple of years, than a bell-shaped hazard results.

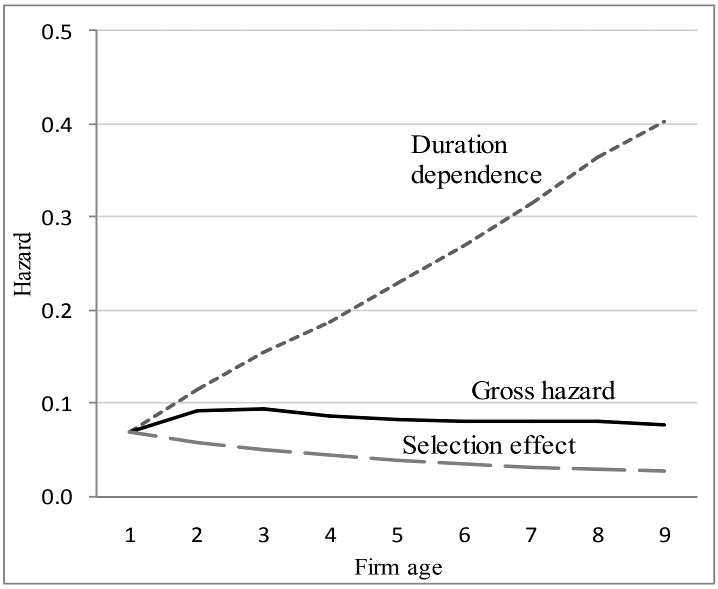

Figure 7 - Gross hazard recomposed

6 Conclusion

Differences between firms creates a selection process. Some of these differences, such as what industry the firm enters, are known to the firm founders before entry, while other differences, such as the firm's productivity, demand and the prices it will face, are gradually learned after entry. I can control for some of the differences observed in the first year after entry, but not all. Controlling for observed differences between firms tilts the hazard function upward compared to the observed exit rate.

It seems plausible that there are significant differences between firms that I am unable to control for. This implies that there is still a selection effect, driven by low quality firms exiting, on average, earlier than high-quality firms. When modelling this remaining heterogeneity explicitly, I get an upward sloping hazard rate. This is the estimated duration dependence, namely what is left of the effect of age on the exit hazard when the selection process has been filtered out.

This duration dependence can be interpreted in terms of vintage theories and active learning. The contribution from the effect of active learning should be a decreasing exit hazard in firm age. Vintage theories predict that the exit hazard increases with the age of capital. Vintage theories have also been used to explain the increasing exit hazard in age in long run, such as in Dale-Olsen (2005) for firms up to 40 years of age.

The results in this paper indicate that the total exit hazard of a firm, given the initial quality of the firm, increases with firm age even in the short run, e.g., a period of up to nine years after firm entry. Positive effects from learning and experience do not seem to be large enough to compensate for the negative effects of ageing (primarily thought of as the effect of old capital equipment in the theoretical literature). However, the duration dependence of the *total exit hazard* obscures important differences between different types of exit. When looking at bankruptcies and liquidations separately, I find that bankruptcies do *not* exhibit a positive duration dependence. Given initial firm quality, the risk of bankruptcy is actually decreasing or at worst constant in firm age. This finding is consistent with the idea that active learning improves the quality of the firm and hence reduces the exit hazard as more experience is accumulated by the firm.

Another important difference between bankruptcy and liquidation is that firm specific parameters are associated with larger effects on bankruptcy risk then on liquidation risk. This includes type of entry, labour productivity, debt-to-equity and the age, education and sex of employees. Firms that go bankrupt appear to be a more uniform group, more easily characterized even at point of entry. The firms with high debt or negative equity, low labour productivity and young, male employees with low education have, on average, a higher risk of bankruptcy.

The magnitude of the duration dependence found in this paper may be exaggerated. But there are good reasons to believe that the true duration dependence of firm liquidation is in fact upward sloping even in the first years. Even the observed liquidation hazard is upward sloping for the first six years. The more differences between firms we do account for, the less negative the slope of the hazard function will become.

It should also be mentioned that the standard errors in these models underestimate the real uncertainty as they ignore the uncertainty arising from model specification errors. One source of model specification error can be the proportionality assumption. For example, if one of the covariates have the effect of altering the time path of the exit hazard, then the estimates are at best representing a crude average relationship. This could be the case: As we have seen, several of the covariates included in this analysis have opposite effects for different types of exit, and the duration dependencies of bankruptcy and liquidation are not at all similar.

Salvanes and Tveteras (2004) found that the risk of plant exit is decreasing in plant age and increasing in capital age, and did not attempt to separate between selection and active learning. The omission of capital age in my analysis might affect the estimates in several ways. The estimated time path of the exit hazard captures some of the effect from the age of capital, if capital age and firm age are in general related. The duration dependence, which we interpret as an effect of firm age, will then include the effect of increasing firm exit due to higher capital age. The omission of capital age is probably also one of the sources of unobserved heterogeneity between firms. The vintage effect is thus likely to be part of the explanation for the positive duration dependence for liquidation hazard found in this analysis.

If we allow for a broad definition of capital in a vintage theory setting, then I believe more of the positive duration dependence of liquidation hazard can be understood in terms of the idea that old capital can be bad for productivity. The sectors with a relatively high turnover rate, such as retail, cafes and restaurants, are naturally overrepresented in an analysis of entering firms, because a larger proportion of firms in such sectors will be new in any given year. The sectors mentioned are characterized by relatively low sunk costs, there is not much machine capital but rather investment in real estate and equipment which can easily be sold and reused by a new firm. For these firms, their name and concept can be important assets. If their popularity declines, the owner might decide to change the name, change the concept, and maybe even redecorate the interior. The old firm will then exit, and a new firm, with new name and new organizational number, will enter with the old plant. In other cases, the owner might prefer to sell the plant and try again somewhere else. Such firm exits will not always be linked to low productivity, high debt or any other easily identified firm characteristic, though in "trend-affected" sectors they do reflect something similar to the vintage effect.

The analysis identifies some of the key components of the observed exit rate of firms in the sample. On the one hand, it represents a mixture of different types of firm exits. The introduction of competing risks had a large impact on our understanding of firm exit, and the findings clearly supports a distinction between different types of exit. On the other hand, the observed exit hazard is composed of both selection effects and a duration component. The selection effect is caused by the observed and unobserved differences between firms, and many of the firm characteristics I have identified in the sample data can contribute to the understanding of why some firms have a higher exit hazard than other. The duration dependence of the exit hazard expresses how the age of the individual firm is related to the probability of exit in our data. When treating all exits together, I found a positive duration dependency. But through the grouping of exits in a competing risks setting, it became clear that the average hides substantial differences between different types of firm exits. The risk of bankruptcy has no significant duration dependency, it does not change with firm age, while the probability of liquidation is increasing as the firms get older.

References

Abbring, J. H, and G. J Van Den Berg (2007): "The unobserved heterogeneity distribution in duration analysis", Biometrika 94, no. 1: 87-99.

Blossfeld, H. P, K. Golsch, and G. Rohwer (2007): Event history analysis with Stata. Lawrence Erlbaum. London

Caves, R. E. (1998): "Industrial organization and new findings on the turnover and mobility of firms", Journal of economic literature 36, no. 4: 1947–1982.

Dale-Olsen, H. (2005): "Etablering og nedlegging av bedrifter, bedrifters størrelse og levetid". In Nytt arbeidsliv: Medvirkning, inkludering og belønning, ed. H. Torp. 250-271. Gyldendal Norsk Forlag, Oslo.

Eckel, C. C. and P. J. Grossman (2003): "Forecasting risk attitudes: An experimental study of actual and forecast risk attitudes of women and men", Virginia Tech Department of Economics Working Paper.

Eurostat (1996): NACE Rev. 1, Statistical Classification of Economic Activities in the European Community. Luxembourg: Eurostat. http://www.chass.utoronto.ca/datalib/other/E0032_en.pdf.

Foster, L., J. Haltiwanger, and C. J. Krizan (2006): "Market selection, reallocation, and restructuring in the US retail trade sector in the 1990s", The Review of Economics and Statistics 88, no. 4: 748–758.

Gaure, S., K. Røed, and T. Zhang (2007): "Time and causality: A Monte Carlo assessment of the timing-of-events approach", Journal of Econometrics 141, no. 2: 1159–1195.

Greene, W. H. (2003): Econometric analysis. 6th ed. Prentice Hall Upper Saddle River, New Jersey.

Huse, M. K. (2009): "Effekt av nedgangstider på vekst i produktivitet", Søkelys på arbeidslivet 3/09: 285-298.

Jenkins, S. P. (1997): "Estimation of discrete time (grouped duration data) proportional hazards models: pgmhaz", Stata Technical Bulletin Reprints: 109–121. Http://www.ecn.ulaval.ca/fileadmin/documents/Proferreurs/Guy_Lacroix/Cours/Lyon/Stata/STB-39-pgmhaz.pdf

Jianakoplos, N. A. and A. Bernasek (1998): "Are women more risk averse?", Economic Inquiry 36, no. 4:620-630.

Jovanovic, B. (1982): "Selection and the Evolution of Industry", Econometrica: Journal of the Econometric Society: 649–670.

Kennedy, P. (2008): A guide to econometrics. 6th ed. Blackwell Publishing.

Regnskapsloven (1998). Lov om årsregnskap m.v. (Regnskapsloven) (The Norwegian

Accounting Act). http://www.lovdata.no/all/nl-19980717-056.html, accessed Jan 16th 2010.

Røed, K., and L. Westlie (2007): "Unemployment insurance in welfare states: Soft Constraints and Mild Sanctions", IZA Discussion Paper no. 2877. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=999378.

Salvanes, K. G, and R. Tveteras (2004): "Plant exit, vintage capital and the business cycle", Journal of Industrial economics 52, no. 2: 255–276.

Schubert, R., M. Brown, M. Gysler and H. Brachinger (1999): "Financial decision making: Are women really more Risk-Averse?", The American Economic Review 89, no. 2: 381-385.

SSB (a). Standard for næringsgruppering (SN2002). http://www3.ssb.no/stabas/ItemsFrames.asp?ID=5556001&Language=nb&Type=hierarchy&VersionLevel=classversion&MenuChoice=Text, accessed Jan 16th 2010.

——(b). Statistikkbanken: Konsumprisindeks. http://statbank.ssb.no/statistikkbanken/Default_FR.asp?PXSid=0&nvl=true&PLanguage=0&tilside=selecttable/hovedtabellHjem.asp&KortnavnWeb=kpi, accessed Jan 16th 2010.

——(c). Årlig nasjonalregnskap 1970-2008. http://www.ssb.no/emner/09/01/nr/tab-18.csv, accessed Jan 16th 2010.

Appendix

Sector code	2-digit NACE code	Sector
a	1-2	Agriculture, hunting and forestry
b	5	Fishing
c	10-14	Mining and quarrying
d	15-37	Manufacturing
e	40-41	Electricity, gas and water supply
f	45	Construction
g	50-52	Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods
h	55	Hotels and restaurants
i	60-64	Transport, storage and communication
j	65-67	Financial intermediation
k	70-74	Real estate, renting and business activities
l	75	Public administration and defence; compulsory social security
m	80	Education
n	85	Health and social work
o	90-93	Other community, social and personal service activities
p	95	Private households with employed persons
q	99	Extra-teretorial organizations and bodies
z	00	unknown

Table 7 - Sector definitions

		Model 1		Model 2		Model 3	
		Basic		Gamma		Discrete	
Variable	Dummy	β	std. β	β	std. β	β	std. β
Sector (Reference sector g)	a	-0.13698	0.11394	-0.19169	0.16562	-0.19428	0.17150
	b	0.04091	0.10194	0.00170	0.15494	0.00718	0.17214
	c	-0.17834	0.16820	-0.23466	0.23732	-0.25553	0.25082
	d	-0.05384	0.03535	-0.03234	0.05382	-0.02794	0.05613
	e	0.09209	0.16814	0.23414	0.24998	0.26231	0.25883
	f	-0.13629	0.03484	-0.15532	0.05092	-0.16431	0.05308
	h	0.24898	0.03356	0.42146	0.05580	0.44427	0.05834
	i	-0.09778	0.04317	-0.09445	0.06498	-0.10633	0.06700
	j	-0.44699	0.11770	-0.54393	0.16539	-0.56740	0.16341
	k	-0.14336	0.02702	-0.18470	0.04099	-0.19986	0.04358
	l	0.69011	0.71282	0.96861	1.50767	1.65132	4.58426
	m	-0.04111	0.10376	-0.09591	0.15332	-0.07704	0.16889
	n	-0.46420	0.06923	-0.59556	0.09413	-0.63153	0.10030
	o	-0.40208	0.05263	-0.59086	0.07737	-0.61472	0.08064
	p	0.83142	1.00293	0.97888	1.38836	0.90566	2.25617
	z	1.03212	0.11844	1.22357	0.22619	1.49367	0.36665

Table 8 - Estimation results models 1, 2 and 3: Sector

		Model 4		Model 5				Model 6			
		All exits		Full exit		Half exit		Bankruptcy		Liquidation	
Variable	Dummy	β	std. β	β	std. β	β	std. β	β	std. β	β	std. β
NACE, 2-digit (Reference 52)	1	-0.32555	0.20182	-0.11604	0.21385	-1.69216	0.53814	-0.35638	0.28068	-0.29156	0.26171
	2	-0.53593	0.32701	-0.42593	0.35121	-1.17892	0.78672	-0.69008	0.45215	-0.25025	0.42197
	5	-0.26425	0.17377	-0.12206	0.18220	-1.02242	0.38409	-0.80276	0.26825	0.26183	0.20303
	11	0.00685	0.39848	0.13982	0.41766	-0.78211	0.89681	-1.66760	1.14682	0.82586	0.43105
	13	-0.40886	1.90906	-0.30328	2.13004	0.00000	0.00000	1.25259	3.66563	0.00000	0.00000
	14	-0.76626	0.33330	-0.64125	0.37032	-1.56243	0.77937	-0.85057	0.49254	-0.50953	0.39017
	15	-0.10366	0.14524	0.06219	0.15392	-0.60438	0.25637	-0.17502	0.20850	-0.00679	0.18039
	16	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	17	-0.26111	0.32178	-0.49033	0.40022	0.05561	0.38216	0.18743	0.43269	-0.31842	0.38083
	18	0.10441	0.39323	0.37084	0.40276	-1.59487	1.14903	-0.11935	0.51937	0.12029	0.49603
	19	0.29759	1.19368	0.13402	1.31221	0.89682	2.84707	0.43601	1.89922	-0.14774	1.76796
	20	-0.15525	0.16987	0.03416	0.17999	-1.01444	0.40560	-0.29258	0.25200	-0.11960	0.21003
	21	-0.21233	0.55435	-0.16232	0.61983	-0.59388	0.95615	0.38763	0.75568	-0.67820	0.81626
	22	-0.29110	0.11838	-0.08894	0.12411	-1.31028	0.27313	-0.34719	0.17272	-0.13006	0.14508
	23	2.02634	1.43589	2.31015	1.52524	0.00000	0.00000	2.90142	2.03003	0.00000	0.00000
	24	-0.02452	0.40442	0.23650	0.42065	0.00000	0.00000	0.09334	0.53376	-0.29497	0.54121
	25	-0.67236	0.35706	-0.39739	0.36887	-2.51036	1.09104	-1.19927	0.57257	-0.13693	0.40315
	26	-0.32588	0.26364	-0.15829	0.28285	-1.43172	0.62799	-0.22367	0.35835	-0.26043	0.32737
	27	0.30801	0.45316	0.31256	0.50333	0.24684	0.63400	0.50411	0.66850	0.03871	0.52982
	28	-0.31807	0.14258	-0.06857	0.15358	-1.63614	0.38032	0.04672	0.19351	-0.62501	0.19885
	29	-0.27929	0.13916	-0.07489	0.14791	-1.33774	0.33604	-0.15053	0.19750	-0.37793	0.17987
	30	-0.49878	0.92869	-0.33547	0.92410	0.00000	0.00000	0.19491	1.27665	-1.28169	1.78530
	31	-0.32813	0.24976	-0.12937	0.27638	-1.38021	0.61570	-0.43659	0.38799	-0.13422	0.29765
	32	-0.62442	0.52711	-0.56796	0.55574	-1.03449	0.89414	-0.99272	0.82828	-0.29060	0.59345
	33	-0.53582	0.25758	-0.27730	0.27226	-2.38544	1.02858	-1.05084	0.43477	-0.21154	0.31281
	34	-1.24935	0.61020	-0.89570	0.65370	0.00000	0.00000	-0.34263	0.82086	0.00000	0.00000
	35	-0.39987	0.19654	-0.17235	0.20766	-1.50376	0.45152	-0.15579	0.27814	-0.68905	0.26679
	36	0.02135	0.17871	0.26165	0.18658	-1.10666	0.44777	0.30340	0.24134	-0.32070	0.23590
	37	-0.54430	0.47198	-0.32098	0.50278	-1.76926	1.17075	-1.64632	0.83201	0.09021	0.53618

Table 9 - Estimation results models 4, 5 and 6: NACE 1/2

		Model 4		Model 5				Model 6			
		All exits		Full exit		Half exit		Bankruptcy		Liquidation	
Variable	Dummy	β	std. β	β	std. β	β	std. β	β	std. β	β	std. β
NACE, 2-digit (Reference 52) cont.	40	-0.01554	0.26031	-0.11790	0.31460	-0.02598	0.37737	-2.49348	0.91802	0.75758	0.30166
	41	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	45	-0.39197	0.06060	-0.17205	0.06424	-1.93124	0.16294	-0.17238	0.08412	-0.59719	0.07999
	50	-0.37162	0.07814	-0.28459	0.08439	-0.72992	0.13855	-0.35799	0.10942	-0.33331	0.10083
	51	-0.54376	0.06429	-0.37521	0.06774	-1.59803	0.15177	-0.50902	0.09205	-0.45704	0.08051
	55	0.26758	0.06068	0.31428	0.06595	0.11735	0.08905	0.48107	0.08318	-0.09126	0.08064
	60	-0.32497	0.09394	-0.13382	0.09911	-1.45689	0.24090	-0.35997	0.13220	-0.34156	0.12285
	61	-0.28118	0.18229	0.02275	0.19063	-3.31643	1.03783	-0.24694	0.27819	-0.25968	0.22852
	62	-0.38114	1.33130	-0.25598	1.44926	0.00000	0.00000	-0.02578	1.65634	-0.35095	1.64014
	63	-0.52031	0.12619	-0.37277	0.13381	-1.32408	0.28786	-0.60513	0.19043	-0.38165	0.15527
	64	-0.05322	0.18102	0.19590	0.19147	-1.46665	0.55907	-0.14465	0.27773	0.00264	0.22863
	65	-1.77437	0.41606	-1.68426	0.42870	-2.50528	1.07983	0.00000	0.00000	-0.88678	0.42725
	66	-1.73432	0.90511	-2.06420	1.02243	-2.20182	2.08246	0.00000	0.00000	-1.63553	1.01851
	67	-0.56454	0.18218	-0.32204	0.18930	-2.85106	1.09620	-0.85357	0.29682	-0.31076	0.22026
	70	-1.07773	0.08915	-0.98490	0.09373	-1.60418	0.20186	-1.36679	0.14678	-0.67718	0.10703
	71	-0.45068	0.17388	-0.28249	0.18747	-1.33628	0.43964	-1.41495	0.30372	0.30123	0.20915
	72	-0.21257	0.07858	-0.01625	0.08319	-1.43679	0.18685	-0.55539	0.12292	0.10172	0.09369
	73	-1.29130	0.57334	-1.23855	0.61656	-1.79949	1.28364	-1.00394	0.89852	-0.87369	0.60244
	74	-0.36119	0.05721	-0.16904	0.06084	-1.68356	0.13538	-0.60142	0.08874	-0.10193	0.06983
	75	1.51609	3.97578	1.88054	4.16850	0.00000	0.00000	0.00000	0.00000	1.09856	4.11891
	80	-0.28531	0.16859	-0.10576	0.17770	-1.29201	0.40438	-0.44533	0.25360	-0.12831	0.20602
	85	-0.82269	0.10370	-0.66601	0.11106	-1.68350	0.22433	-1.70586	0.21020	-0.30025	0.11807
	88	0.91849	0.86427	1.17642	0.91392	0.00000	0.00000	0.76593	1.00045	1.94296	0.94282
	90	-1.28744	0.32950	-1.44182	0.38793	-0.98679	0.48180	-2.36839	0.65031	-0.51433	0.37506
	91	0.22025	0.51202	0.51200	0.53523	0.00000	0.00000	-0.55716	0.83570	0.53804	0.72791
	92	-0.77512	0.12092	-0.65257	0.12752	-1.35859	0.24899	-0.94866	0.18239	-0.48352	0.14889
	93	-0.79790	0.11213	-0.75427	0.12195	-1.09801	0.18981	-1.14263	0.17921	-0.33348	0.13516
	95	0.75378	2.16048	0.91399	2.37119	0.00000	0.00000	1.41184	3.14633	0.00000	0.00000
	98	1.34118	0.34859	1.52774	0.34758	-1.11302	1.10779	0.93744	0.48088	1.43293	0.40473

Table 10 - Estimation results models 4, 5 and 6: NACE 2/2

		Model 4		Model 5				Model 6			
		All exits		Full exit		Half exit		Bankruptcy		Liquidation	
Variable	Dummy	β	std. β	β	std. β	β	std. β	β	std. β	β	std. β
County (Reference Oslo)	Østfold	-0.32243	0.07580	-0.35921	0.08007	-0.21431	0.15288	-0.17547	0.11018	-0.37680	0.09522
	Akershus	-0.20457	0.05735	-0.21919	0.06073	-0.17043	0.11984	-0.25983	0.08591	-0.14557	0.06954
	Hedmark	-0.21571	0.08787	-0.22580	0.09289	-0.20238	0.17805	0.04171	0.12359	-0.49553	0.11562
	Oppland	-0.13716	0.08693	-0.14036	0.09278	-0.12998	0.16662	-0.01257	0.12528	-0.25224	0.11029
	Buskerud	-0.10972	0.07141	-0.11880	0.07579	-0.07303	0.14636	-0.11943	0.10525	-0.13422	0.08894
	Vestfold	-0.24888	0.07327	-0.28582	0.07823	-0.13350	0.14814	-0.12915	0.10719	-0.33618	0.09124
	Telemark	-0.12888	0.08966	-0.17788	0.09542	0.06200	0.17023	-0.12096	0.13005	-0.09999	0.10949
	Aust-Agder	-0.21585	0.10645	-0.30619	0.11475	0.07075	0.19185	-0.18101	0.15834	-0.17865	0.12838
	Vest-Agder	-0.28383	0.08746	-0.33848	0.09453	-0.09320	0.15752	-0.43385	0.13704	-0.09906	0.10263
	Rogaland	-0.13237	0.06095	-0.14168	0.06466	-0.15235	0.12211	-0.17458	0.09249	-0.09766	0.07371
	Hordaland	-0.27667	0.06141	-0.29169	0.06529	-0.27532	0.12753	-0.03082	0.08924	-0.54948	0.07754
	Sogn og Fjordane	-0.27112	0.10423	-0.27475	0.11085	-0.34652	0.21047	0.07299	0.15180	-0.67338	0.13498
	Møre og Romsdal	-0.20962	0.07683	-0.21602	0.08160	-0.25446	0.15568	0.07406	0.10782	-0.46094	0.09863
	Sør-Trøndelag	-0.15018	0.06963	-0.19074	0.07436	0.04506	0.13603	-0.07605	0.10374	-0.25522	0.08628
	Nord-Trøndelag	-0.19247	0.09954	-0.17144	0.10461	-0.34628	0.21680	0.05400	0.14241	-0.44479	0.12626
	Nordland	0.14699	0.07222	0.18384	0.07652	-0.08776	0.14598	0.56787	0.10509	-0.52163	0.09790
	Troms	-0.20518	0.08865	-0.20400	0.09398	-0.28020	0.17769	0.06471	0.12482	-0.38201	0.11085
	Finnmark	0.00405	0.10770	0.07667	0.11564	-0.56633	0.24122	0.42268	0.14584	-0.46230	0.14313
	Svalbard	-0.74469	0.87463	-0.83294	0.92130	-0.59409	1.24272	-2.18450	1.61530	0.26496	0.77805
	Unknown	1.89077	0.22578	1.86879	0.24035	2.22055	0.37542	2.12647	0.32503	1.57254	0.25646

Table 11 - Estimation results models 4, 5 and 6: County