SNF REPORT No. 04/10

Determinants of Long-Term Bank Relationships

An empirical study of the Norwegian bank market

by

Robert Christensen and Sindre Johansen

SNF Project No. 7556
Analyse av kapitaltilgang og kapitalbruk – Foretaksfinansiering og handlemønstre i verdipapirmarkedet (The Finance Market Fund, Project 189007)

The project is supported by The Finance Market Fund

INSTITUTE FOR RESEARCH IN ECONOMICS AND BUSINESS ADMINISTRATION
BERGEN, April 2010

ISBN 978-82-491-0697-4 Printed version
ISBN 978-82-491-0698-1 Online version
ISSN 0803-4036

Preface

This thesis was written as a part of the Master of Science programme at The Norwegian School of Economics and Business Administration, NHH. Working with the modern bank relationship literature, we have gained insight into determinants of the duration of firm-bank relationships. The goal of this paper was to conduct an empirical analysis of potential determinants on the Norwegian market. While there are several papers written about how strong bank relationships can be valuable for firms, only a limited number of studies focus on the length of the relationship.

We recognize that our sample is right-censored which is a common challenge related to duration studies. This challenge can bias inferences about the length of bank relationships. Despite the challenge this study complements the existing literature, and the main contribution of the paper is the study of new explanatory variables on a relatively large data sample.

We would like to thank our thesis advisors, Dr.oecon Gorm A. Grønnevet and Dr.oecon Aksel Mjøs, for valuable comments and firm guidance during the process. Our gratitude is also extended to the Ph.D. candidate Ove R. Hetland for fruitful discussions, and to the Ministry of Finance for providing the data that form the backbone of this empirical analysis.

Bergen, 20 December 2009

Robert Christensen

Sindre Johansen

Contents

Abstract

Preface

1. Introduction 1

2. Literature review 3

Empirical implications 8

3. Data and Sample Selection 9

3.1 Data collection 9

3.2 Sample 10

3.3 Challenges with the data 17

3.4 Determinants of relationship duration 21

4. Methodology and Econometric specification 28

4.1 Survival and duration analysis 28

4.2 Comparison of the survival models 35

4.3 Assumptions about the explanatory variables 35

5. Empirical results 37

5.1 Cox partial likelihood estimates 38

5.2 Results using restricted baseline hazard models 45

5.3 Robustness tests 50

6. Discussion and conclusion 53

7. References 57

8. Appendices 62

8.1 Data overview 63

8.2 Econometric definitions 71

8.3 Results 72

List of Figures

Figure 1 Distributions of firms by the number of bank relationships 12
Figure 2 Distribution of relationships, by type of bank 14
Figure 3 Distribution of relationships, by type of account 15
Figure 4 Distribution of relationships, by bank owner nationality 16
Figure 5 Distribution of relationships, by firm age 16
Figure 6 Censoring..... 28
Figure 7 Kaplan-Meier survival estimate (smoothed) 37

List of Tables

Table 1 Evidence on Bank relationships: Duration 4
Table 2 Evidence on Bank relationships: Number and concentration 7
Table 3 Annual overview of the sample 11
Table 4 Distribution of observed duration of bank relationships..... 13
Table 5 Characteristics of which the positive impact on duration may be underestimated .. 21
Table 6 Descriptive statistics for explanatory variables 25
Table 7 Correlation matrix. 27
Table 8 Overview of explanatory variables included in our survival analysis 39
Table 9 Partial likelihood estimates of proportional hazard model (Cox)..... 40
Table 10 Parametric estimation of proportional hazard model (Exponential and Weibull) ... 47
Table 11 Robustness tests. Estimations of the Weibull model..... 50

1. Introduction

A general definition of a bank relationship is *"the connection between a bank and a firm that goes beyond the execution of simple, anonymous, financial transactions*[1].The recent literature on financial intermediation primarily focuses on the role of banks as relationship lenders. As lenders, banks are important to firms because they are often a firm's main source of external financing. This is indeed the case for the Norwegian market where, in 2008, almost 80% of all commercial debt was covered by banks.[2] Further, a central aspect in the literature on bank relationships is information asymmetries. Diamond (1984) argues that in lasting relationships, the bank is able to monitor the performance of the firm. Along those lines, the bank can mitigate the information problems since the extent of private information increases over time.

By reviewing the literature it becomes clear that from a firm's perspective, there are benefits and drawbacks related to maintaining long-term bank relationships[3]. In short, a strong bank relationship is in general thought of as valuable since it reduces costs (e.g. a more favorable rate on the loan contracts) and increases the availability of new funding. A firm with a good reputation as a trustworthy borrower (e.g. the firm has a good payment record), can also achieve better contract conditions. On the other hand, the cost of a strong bank relationship is mainly related to the so called holdup problem; a bank can take advantage of the accumulated private firm information it has collected over the course of the relationship by "locking-in" the firm and extracting monopoly benefits.

Based on the benefits and drawbacks of a strong bank relationship, Ongena and Smith (1998b) argue that the duration of the relationship itself may influence a firm's decision to end the relationship. In other words, the duration of the relationship should reflect the net impact of the factors mentioned above. Their logic can be summed up in two statements:

(a) *If the value of a relationship tends to increase through time, or, if firms become locked into specific banks, the likelihood of terminating a relationship should decline over time*

[1] Ongena and Smith (1998a): "Banking relationships: a review".
[2] Statistical Yearbook of Norway 2008 (Statistics Norway, SSB).
[3] See Bhattacharya and Thakor (1993) for a review.

(b) If the value of a relationship tends to decrease over time, and if the switching costs are relatively low, the likelihood of terminating a relationship should increase through time

However, time cannot be the only factor influencing the value of bank relationships. This can be taken into account by studying firm-specific characteristics. Ongena and Smith (1998b) find that firms are more likely to end a relationship as the relationship matures, supporting the statement *(b)* above; the value of the relationship decreases through time. Further, the study shows that small, profitable and highly leveraged firms maintain the shortest relationships. Firms taking up loans in more than one bank are also found to maintain shorter relationships.

In our study we utilize data obtained from the Ministry of Finance and the Brønnøysund Register (Brønnøysundregistrene)[4] to study the value of long-term bank relationships in the period from 1998 to 2008. Like Ongena and Smith we focus on the willingness and ability of firms to end bank relationships. The data allow us to derive information on bank relationships from a time series of annual data, which enables us to observe the beginning, development and end of most of the existing relationships in our sample. In our study we merge relationship data with firm- and bank-specific data, and use hazard function estimators to draw conclusions on the factors that affect relationship length. By focusing only on lending relationships that we know started in 1998, we avoid left-censoring problems. However, we carefully recognize that our sample is *right-censored;* 16% of the relationships in our sample are still ongoing in our last year of data. These relationships continue after 2008, censoring our maximum observable relationship duration to 11 years. Due to the fact that our study is only corrected for left-censoring, it is possible that our inferences about the length of the relationships are biased. Another (potential) weakness of our study is that the effects of mergers and acquisitions are only partly taken into account. We discuss this issue by conducting a case study.

We find that firms become more likely to end a bank relationship as the relationship matures. As mentioned this suggests a decreasing value of relationships over time, that switching costs are not prohibitively large, and that firms try to avoid holdup costs.

[4] A government body under the Norwegian Ministry of Trade and Industry, and consists of several different national computerized registers.

Further, we find the median duration of a bank relationship started in 1998 to be 4 years. By testing for various determinants of ending a bank relationship, we find that the likelihood of ending a bank relationship decreases when a firm maintains a relationship with a savings bank and holds both deposits and loans. We also find that the length of the relationship is longer when a firm has a high tangibility ratio, has high creditor concentration and is highly profitable. Firms are more likely to end a relationship early when it is with a Norwegian (owned) bank and when a firm holds multiple bank relationships.

It should be mentioned that all inferences about the ending of the relationships are drawn from what is revealed through the observed terminations. In other words, we do not have direct information about the real cause of the termination (nor of the start) of the relationships. Hence, it is for example impossible for us to measure the value added/lost due to long-term relationships from the firm's perspective. Nevertheless, the main contribution of our study to the existing literature is that the unique dataset used enables us to study new explanatory variables on a relatively large sample.

The rest of the paper is organized as follows: The next section provides an overview of the theoretical literature relevant to our study. Section 3 describes the data used in our analysis. Section 4 contains information about the econometric specifications used for analyzing duration data, including an introduction to both semiparametric and parametric proportional hazard models. Section 5 contains empirical results including robustness tests, and finally section 6 concludes.

2. Literature review

Through the progression of a relationship, a bank learns more about a firm's ability to meet future obligations, both from past payment history and through other services offered by the bank. The importance of a relationship will depend on the length or duration of the interaction between the customer and the bank. Rajan (1997) explains that a long term interaction between two parties is mutually beneficial, and that relationships may evolve in situations where explicit contracts are inadequate. The duration seems to have a positive influence on the availability of credit for firms, and firms in longer bank relationships are less

likely to pledge collateral against the loan (Ongena and Smith, 1998a). The average duration of a relationship depends on which study or data set we look at, as shown in Table 1. Ongena and Smith (1998b) estimate the average duration of a bank relationship in Norway to be six years. They find that long-term relationships are more likely to be terminated than shorter relationships. This is consistent with the assumptions made in Greenbaum et al. (1989) that the value of a bank relationship decreases with the relationships length due to holdup costs.

Table 1: Evidence on Bank relationships - Duration

Paper	Country	Sample Year(s)	Sample Size	Average (Median) Firm Size	Average (Median) Duration, years
Cole (1998)	US	1993	5,356	Book assets: USD 1.63m	7.03
Blackwell and Winters (1997)	US	1988	174	Book assets: USD 13.5m	9.01
Petersen and Rajan (1995)	US	1987	3,404	Book assets: USD 1.05m (0.3m) Employees: 26 (5)	11.3
Angelini et al. (1998)	Italy	1995	1,858	Employees: 10.3	14.0
Harhoff and Körting (1998)	Germany	1997	994	Employees: ± 40 (10)	± 12
Elsas and Krahnen (1997)	Germany	1992-1996	125 / year	Turnover: (30-150)	22.2
Kim et al. (2001)	Norway	1988-1996	177 / year	Book assets: USD 4m	13.5
Ongena and Smith (1998)	Norway	1979-1995	111 / year	Market equity: USD 150m	(15.8 - 18.1)
Zineldin (1995)	Sweden	1994	179	Employees: (<49)	(>5)
Sjögren (1994)	Sweden	1916-1947	50	Largest firms	>20 (5-29)
Degryse and Van Cayseele (1998)	Belgium	1997	17,776 loans	Employees: (1)	7.82
Horiuchi et al. (1988)	Japan	1962-1972	479	Largest firms	(21)
		1972-1983	668		(30)

Notes: Sample size is the number of firms (unless indicated otherwise). Average (median) firm size is the size of the firms in millions of US$ in the last year of the sample or if indicated the number of employees. Average (median) duration is the duration of bank relationships in years. Source: Ongena and Smith (1998a)

Holdup costs

Firm-bank relationships are particularly important in bank-centered financial systems found in continental-Europe and in Scandinavia. Banks supply almost 80% of all commercial lending in Norway.[5] Strong bank relationships are often considered valuable for the firm, and maintaining a relationship often means that the customer and bank are willing to make temporary sacrifices in favor of obtaining future benefits. However, as a bank acquires private information during the course of a relationship, this could easily "lock-in" the customer, and the bank may be tempted to charge monopoly rents. Among others[6], von Thadden (1995) analyses the value of banking relationships in the presence of such holdup costs; the firm must pay a "lemon's premium" if it approaches another bank, as this bank will question the firm's decision not to seek funding at its original lender. As previously

[5] Statistical Yearbook of Norway 2008 (Statistics Norway, SSB).
[6] Greenbaum et al. (1989), Sharpe (1990), Rajan (1992), Petersen and Rajan (1995).

mentioned, Greenbaum et al (1989) argue that the value of the holdup rents decline with the length of the relationship; as the bank gathers more information from its customers through time, uncertainty about a given customer's cash flow declines, mitigating the holdup problem. The holdup problem is not favorable for a firm. A seemingly simple solution to the problem is for a firm to establish more than one bank relationship and then let the banks compete away the monopoly pricing as more symmetric information is obtained. The issue of multiple-bank relationships is discussed further later in the section.

Many factors, both firm specific and bank specific, influence the duration of a relationship . We review the factors we find most important and present previous studies' findings and conclusions to help illuminate and justify our further analysis.

Firm size and age

Petersen and Rajan (1994) find that the firm's age is the most important determinant in explaining the variation in lending rates, where older firms get more favorable conditions. This implies that old firms may have longer bank relationships due to more beneficial interest rate margins.

Rajan (1992) finds that small firms are more likely to have problems with private information asymmetry, thus more likely to find bank financing useful. Similarly, Elsas and Krahnen (1998) conclude that small firms are more likely to suffer from assumed higher risk, thus justifying a higher premium on loans. Then again, after a bank loan is initialized Fama (1985) and Diamond (1991) argue that it is primarily small, less prestigious firms that gain most from screening and monitoring services associated with bank loans.

Ongena and Smith (1998b) document that bank relationships have a tendency to be shorter for small, young firms with relatively high leverage. This indicates the same as Farinha and Santos (2000) findings; that small, young and leveraged firms with the greatest need of bank financing, stay in their relationships relatively briefly. In addition, Ongena and Smith (1998b) report that firms with multiple simultaneous relationships terminate a relationship earlier than a firm with only one relationship.

Degryse et al. (2009) show that across Europe, the switching rates of bank relationships are dependent on firm size, and that the rates change sign going from North to South in Europe.

Small firms end a relationship more easily than large firms in Norway, Denmark and Belgium, at the same rate in the United Kingdom and Germany, but at a slower rate in Portugal and Italy. Large firms in Norway prefer a few steady relationships, while in Italy large firms continue to juggle, and drop, many relationships.

Multiple relationships

Ongena and Smith (1998b) find that multiple-bank firms tend to turn over newer relationships and keep one long-term relationship. However, a bank relationship can persist because switching costs are so high that a switch would not be beneficial. Therefore, long-term relationships appear valuable to firms that are unlikely to face credible holdup threats from one monopolistic bank. Farinha and Santos (2000) find that the likelihood of moving to multiple bank relationships is increasing in relationship duration, firm leverage and firm growth potential. The majority of firm's first loan is from a single bank, but shortly afterwards some of the firms start borrowing from several banks. Even small firms that would benefit most from strong bank relationships borrow from multiple banks (Guiso and Minetti (2004)). Houston and James (1996) prove that more than 60% of listed US firms have multiple bank relationships, and Detragiache et al. (1997) claim that the need for multiple relationships emerges through the firm's need to diversify its bank liquidity risk. Then again, Bris and Welch (2005) argue that strong firms choose fewer creditors to signal their confidence in not going bankrupt.

Generally, Farinha and Santos (2000) show that a likely explanation for initiating multiple bank relationships is the bank's unwillingness to increase the lending level because of a firm's past poor performance, rather than the firm's need to protect themselves against the holdup costs one single relationship contains. The literature on the implications of duration on single versus multiple relationships[7] finds that longer-term bank relationships mainly improve the credit availability, but have mixed effect on the interest rate level offered. Petersen and Rajan (1994) find that firms with multiple relationships pay higher interest rates and are more credit restrained than the firms with a single relationship.[8] Petersen and Rajan (1994) and Ongena and Smith (1998b) document empirical evidence that supports the

[7] See Petersen and Rajan (1994), Berger and Udell (1994), Cole (1998), Elsas and Krahnen (1998), Harhoff and Körting (1998), Angelini, Salvo and Ferri (1998), Degryse and Van Cayseele (1998) or D'Auria, Foglia and Reedtz (1999).

[8] D'Auria, Foglia and Reedtz (1999), however, find that firms with more relationships pay lower interest rates.

theory that a single bank is the optimal number of relationships. Both studies estimated the median of bank relationships per firm to be one. However, similarly to the duration of a relationship, the average and optimal number of bank relationships varies across studies and data sets. A summary is given in Table 2 on the next page.

Table 2: Evidence on Bank relationships - Number and concentration

Paper	Country	Sample Year(s)	Sample Size	Average (Median) Firm Size	Average (Median) Number	Concentration
Ongena and Smith (1998)	*Average 20 countries*	*1996*	*1.129*	*Sales: USD 750m*	*5.6*	
	Italy		*70*	*Sales: USD 1,500m*	*15.2*	*35.9*
Detragiache et al. (1997)	Italy	1989-1993	±1,000 / year	Employees: 926 (293)	16.4 (13)	
Pagano, Panetta and Zingales (1998)	Italy	1982-1992	19.274	Employees: 737 (258)	13.9 (11)	
Rossignoli and Chesini (1995)	Italy	1993	1.527		14.8	
Angelini et al. (1998)	Italy	1995	1.858	Employees: 10.3	2.4	
Cesarini (1994)	Italy	1993	263.376	Credit line: < 1 bln. Lira	1.6	
				Credit line: > 500 bln. Lira	33.2	
	Portugal		*43*	*Sales: USD 750m*	*11.5*	*38.1*
Farinha and Santos (2000)	Portugal	1980-1996	54.182		1.00	
	France		*25*	*Sales: USD 1,500m*	*11.3*	*63.6*
	Belgium		*10*	*Sales: USD 3,500m*	*11.1*	*44.4*
	Spain		*68*	*Sales: USD 1,500m*	*9.7*	*50.1*
	Germany		*67*	*Sales: USD 3,500m*	*8.1*	*89.5*
Elsas and Krahnen (1997)	Germany	1992-1996	125 / year	Turnover: (30-150)	6.0 (5.0)	
Harhoff and Körting (1998)	Germany	1997	994	Employees: ± 40 (10)	1.8 (1 or 2)	
	Greece		*41*	*Sales: USD 750m*	*7.4*	*98.3*
	Austria		*37*	*Sales: USD 1,500m*	*5.2*	*61.4*
	Luxembourg		*8*	*Sales: USD 375m*	*5.0*	*17.2*
	Czech Rep.		*59*	*Sales: < USD 100m*	*4.7*	
	Hungary		*44*	*Sales: USD 175m*	*4.0*	
	Finland		*89*	*Sales: USD 750m*	*3.6*	*93.8*
	Switzerland		*39*	*Sales: USD 3,500m*	*3.6*	*79.8*
	Denmark		*51*	*Sales: USD 750m*	*3.5*	*63.7*
	Netherlands		*49*	*Sales: USD 1,500m*	*3.5*	*59.0*
	Poland		*13*	*Sales: USD 175m*	*3.3*	
	Ireland		*67*	*Sales: USD 750m*	*3.2*	*93.6*
	UK		*142*	*Sales: USD 1,500m*	*2.9*	*29.1*
	Sweden		*50*	*Sales: USD 1,500m*	*2.5*	*86.6*
Zineldin (1995)	Sweden	1994	179	Employees: (<49)	(1)	
Berglöf and Sjögren (1995)	Sweden	1984, '90, '93	± 30 / year	Large firms	(1)	
	Norway		*41*	*Sales: USD 750m*	*2.3*	*48.8*
Ongena and Smith (1998)	Norway	1979-1995	111 / year	Market equity: USD 150m	1.4 (1)	
Horiuchi (1993), Horiuchi (1994)	Japan	1990	126 / 309	Employees: < 300 / >300	3.4 / 7.7	28.3
		1992	175 / 189	Employees: < 10 / >10	2.9 (3) / 3.1 (3)	
Petersen and Rajan (1995), Berger and Udell (1995)	US	1987	3.404	Book assets: USD 1.05m (0.3m) Employees: 26 (5)	1.6 (1)	13.3
Houston and James (1996)	US	1980, '85, '90	±250 / year	Market assets: USD 1,502m (112m)	5.22	
Masonson (1992)	US	1991	1.123	Sales: > USD 500m	7.0	

Notes: Sample size is the number of firms (unless indicated otherwise). Average (median) firm size is the size of the firms in millions of US$ in the last year of the sample, or, if indicated, the number of employees. Average (median) number is the number of bank relationships. The concentration ratio is the percentage of total banking system assets accounted for by the largest three banks in the country in 1993. All data in italic depicted from Barth et al. (1997) Source: Ongena and Smith (1998a).

Creditor concentration

Obviously, a one-bank firm has only one bank financing source, but Petersen and Rajan (1994) show that even firms with multiple lending sources tend to concentrate their loans around one source. This is particularly evident for small firms, while the pattern becomes less evident as the firm size grows. Further, Ongena et al (2007) find that higher quality firms and firms with more liquid assets choose more concentrated borrowing, which is consistent with the findings of Bris and Welch (2005) that firm quality has a positive effect on creditor concentration. Similar to Petersen and Rajan (1994), Ongena et al. (2007) find that smaller firms and less leveraged firms have higher concentration in their borrowing.

Bank characteristics

Kim et al. (2004) focus on quality characteristics that banks adopt in order to differentiate themselves from competing banks in order to attract borrowers. Well diversified and well capitalized banks will less likely face large losses and are more able to withstand potential loss. Because of this, firms prefer borrowing from larger banks. Borrowers may believe a larger bank is more likely to be considered as "too big to fail" by the government.

Empirical implications

To summarize, the theoretical literature provides the groundwork for our study of firm-bank relationships through time and the likelihood that a firm may terminate a bank relationship. Ongena and Smith (1998b) find that long-term relationships are more likely to be terminated than shorter relationships, which is consistent with the findings of Greenbaum et al. (1989) that holdup costs contribute to decreasing the value of a bank relationship. Firms are afraid of getting "locked-in", as this may entice banks to charge monopoly rents. Initiating multiple relationships is a possible solution to the holdup problem, but turning over a bank relationship may be restricted by switching costs. Characteristics of the firm are important when analyzing the determinants influencing the relationship duration; Rajan (1992) and Petersen and Rajan (1994) find that a firm's size and age are the most important factors influencing the lending rate. "The bigger the better" is the trend in obtaining favorable rates and thus implies longer relationships, which is consistent with Elsas and Krahnens' (1998)

findings. Degryse et al. (2009) show that size matters when it comes to the turnover of bank relationships across Europe.

Ongena and Smith (1998b) and Farinha and Santos (2000) find that firms which are most likely to benefit from strong relationships (small, young and leveraged firms) actually stay in their relationships relatively briefly. Further, Petersen and Rajan (1994) find that firms with multiple lending sources tend to concentrate their borrowings around one source, which may imply a longer, more stable relationship. Kim et al. (2004) emphasize the characteristics and diversification of bank services as important explanations of the value of a relationship.

3. Data and Sample Selection

3.1 Data collection

Roughly, the data sources can be divided in two groups; (1) relationship data and (2) firm-specific data.

Relationship data

The unique annual data on bank relationships in the period from 1998 to 2008 is obtained from the Norwegian Ministry of Finance. These data are based on information about loan- and deposit relationships between firms and banks/credit institutions, which all Norwegian banks and other credit institutions have to report to the tax authorities on behalf of firms. For all firms, the data contains information about the end-of-year status for every single loan/deposit account. The amount of interests paid during the year is also reported. In addition, for each account there is information about which bank category the bank belongs to. Due to the highly sensitive nature of this information, these data are not publicly available. Hence our duration study on the Norwegian market is the first that is conducted on such a rich data set.

Firm-specific data

Firm-specific data are obtained from the Dun & Bradstreet database[9]. The database is delivered annually to The Norwegians School of Economics and Business Administration

[9] Dun & Bradstreet develops this database on behalf of the *Brønnøysund Registry*

(NHH) and SNF[10]. Dun & Bradstreet collect firms' annual accounting data, such as income statement and the balance sheet. Other firm specific information is also available, i.a. funding year, number of employees and the region in which the company is located. Mjøs and Øksnes (2009) provide a thorough description of the Dun & Bradstreet database.

3.2 Sample

In a study of relationship duration, the analysis may be biased if it is not known when the relationship started or ended.[11] In order to avoid the problem of not knowing when relationships are initiated, we identify all new relationships that started in 1998, and limit our sample to those. Further, since the literature and empirical studies within the banking relationship field focus on the lending side, we identify and remove the relationships that are solely characterized by being a "deposit relationship" throughout the whole period (from 1998 to 2008). Finally since we focus on relationships between firms and banks we also ignore interbank relationships[12].

By making this selection we obtain a sample of 9,476 individual firms, and each firm is identified by its unique organization number[13].

We will now provide a descriptive overview of our selected sample. Since the overview covers our sample only, it is not a general market outline.

During the 11 year sample period 84%, or 7,992 of the 9,476 firms in our sample, terminate one or more bank relationships. Further, already after two years (in 2000) about 50% of the firms had terminated at least one relationship. Table 3 presents an annual overview of the number of firms in our sample, along with the number of firms that have ended one or more relationships.

[10] Institute for Research in Economics and Business Administration (www.snf.no).
[11] This problem is called censoring and is explained in more details in the next section about methodology and econometric specification.
[12] Relationships where a bank/credit institutions lends money to another bank.
[13] In Norway each firm registered in the Central Coordinating Register for Legal Entities is assigned a unique nine digit *organization number* by the *Brønnøysund Register/Ministry of Trade and Industry.*

Table 3: Annual overview of the sample [14]

Year	Firms in Sample	Firms ending bank relationship	% marginal
1998	9,476	1,781	18.8 %
1999	7,889	1,777	22.5 %
2000	6,251	1,090	17.4 %
2001	5,238	1,644	31.4 %
2002	3,650	565	15.5 %
2003	3,108	549	17.7 %
2004	2,580	322	12.5 %
2005	2,269	206	9.1 %
2006	2,071	288	13.9 %
2007	1,789	194	10.8 %
2008	1,600	Unknown	Unknown

This table lists, by year, the total number of unique firms in our sample and the number, and percentage, of unique firms that are ending one or more bank relationships. A bank relationship is defined as a relationship between a firm and a bank. In all of the relationships the bank provides loan(s) to the firm. We identify a firm as ending a relationship when the relationship is no longer reported to the tax authorities. All numbers are obtained from the Norwegian Ministry of Finance. For 2008 the number of firms that end a relationship is unknown due to right-censoring.

When all of a firm's relationships are terminated, the firm leaves the sample[15]. Thus, in the table, the number of firms is declining. We observe that the annual average of firms that end at least one relationship is 17%. Note however that in 2001, this figure almost doubles to 31%, a year characterized as the beginning of a three year economical downturn.[16] This higher rate of termination may be explained by firms defaulting on their loans, combined with banks being more reluctant with regard to new funding (to avoid losses). It is then probable that firms in need of funding switched banks. Another potential reason for the increase in figures in 2001 is new structural changes; international banks, primarily Nordic ones, heavily enter the Norwegian market and this leads to a considerable increase of the bank market competition.[17]

[14] See Table A in the Appendix for the cumulative distribution of when firms are ending relationships
[15] The firm is leaving the sample even if it starts new relationship(s) with other banks.
[16] Norway experienced strong economical growth from 1993 to 1998, after which the growth diminished. Then, the downturn from 2001 to 2003 came in the wake of the IT bubble burst (and 9/11). (Statistics Norway, SSB, 2008).
[17] The Norwegian Savings Banks Association (Sparebankforeningen), 2001

Figure 1 shows the distribution of the sample firms by the number of bank relationships maintained at the beginning of the sample period in 1998 when all the relationships are new.[18]

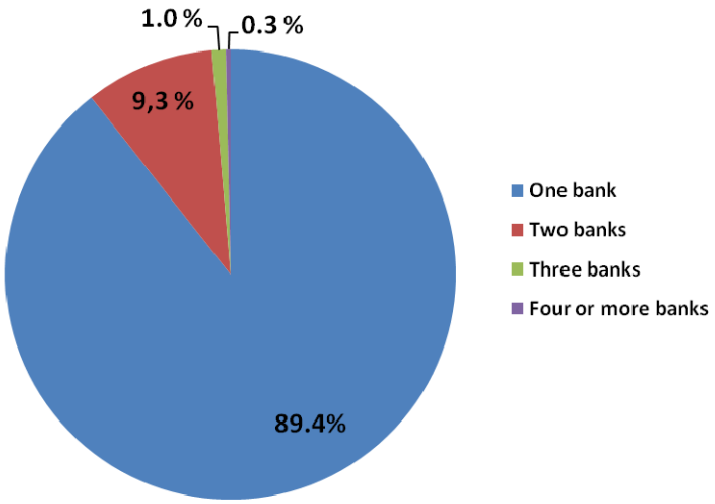


Figure 1: Distributions of firms by the number of bank relationships (1998)

This figure shows the distribution of firms by the number of bank relationships in 1998, when all the relationships in the sample start. A bank relationship is defined as a relationship between a firm and a bank. In all of the relationships the bank provides loan(s) to the firm.

As we can see, in 1998, 89.4 % of the firms have a relationship to only one lending bank, 9.3% have two relationships while only 1.3% of the firms have three or more relationships.[19] The observed average number of relationships is 1.1, and the median equals one. The percentage of firms having multiple relationships is decreasing as firms terminate relationships, and by the end of the sample period, there are 1,600 firms left in the sample, and only 27 (1.7%) maintain more than one relationship.

In the following analysis we define a sample observation as a firm-bank relationship. This is the same as one relationship between a firm and a bank. Following this definition the number of firm-bank relationships is given by the total number of unique firms in the sample

[18] More detailed information can be found in Table B in the appendix.

[19] Note that this is not the same as saying that during an average year 89.4% of the firms maintained a relationship to only one bank. Since we focus on new observations that occurred in 1998, to avoid left-censoring problems, we are not taking into account the fact that the firms may have had other relationship(s) that started before 1997.

in 1998 (9,476), multiplied with the number of relationships that were started by each of these firms in 1998. Our dataset contains 10,130 such observations, justifying the average number of sample observations to be the previously stated 1.1.[20]

By identifying the last year a firm-bank relationship is reported to the tax authorities, we are able to recognize when a relationship is terminated. By the same logic we are able to count the number of consecutive years of each relationship, and we define this to be the duration of a relationship.

Given our definitions and assumptions the distribution of observed duration bank relationships is presented in the following table.

Table 4: Distribution of observed duration of bank relationships

Observed duration, in years	Number of relationships	%, marginal	%, cumulative
1	1,818	17.9 %	17.9 %
2	1,805	17.8 %	35.8 %
3	1,103	10.9 %	46.7 %
4	1,655	16.3 %	63.0 %
5	571	5.6 %	68.6 %
6	552	5.4 %	74.1 %
7	322	3.2 %	77.3 %
8	206	2.0 %	79.3 %
9	290	2.9 %	82.2 %
10	195	1.9 %	84.1 %
≥11	1,613	15.9 %	100.0 %

This table lists the number and the percentage (marginal and cumulative) distributions for the observed relationships, in years. A bank relationship is defined as a relationship between a firm and a bank and in all of the relationships the bank is providing loan(s) to the firm. The total number of firm-bank relationships (10,130) is given by the total number of unique firms in our sample (9,476), multiplied with the number of relationships that were started by each of these firms in 1998. We define a firm as ending a relationship when the firm is no longer reporting the relationship to the tax authorities. The observed duration of a relationship is the number of consecutive years a firm reports the relationship to the tax authorities. All input data is obtained from the Norwegian Ministry of Finance.

As we can see the majority (63 %) of the observations last for less than five years, and already after the first year about 18% of the relationships are terminated. The median duration of an observed relationship is four years. Note also that 15.9% of the observations continued beyond the 11-year cutoff of the sample. The table suggests that the

[20] Note that in the original data set all accounts that are held by a firm with a bank are divided into single account observations. We merge the accounts by summing the firm's total amount of deposits and loans with the bank.

relationships in our sample are short-lived and that the end-of-sample cutoff affects only a rather small proportion of the relationships.[21]

We are now going to look at a breakdown of bank categories represented in the sample. As can be seen in Figure 2, in 1998 40% of the relationships are between a firm and a commercial bank, and 28% are between a firm and a savings bank. The remaining 32% are related to other type of lenders, dominated by state/municipal banks, insurance firms and credit firms.

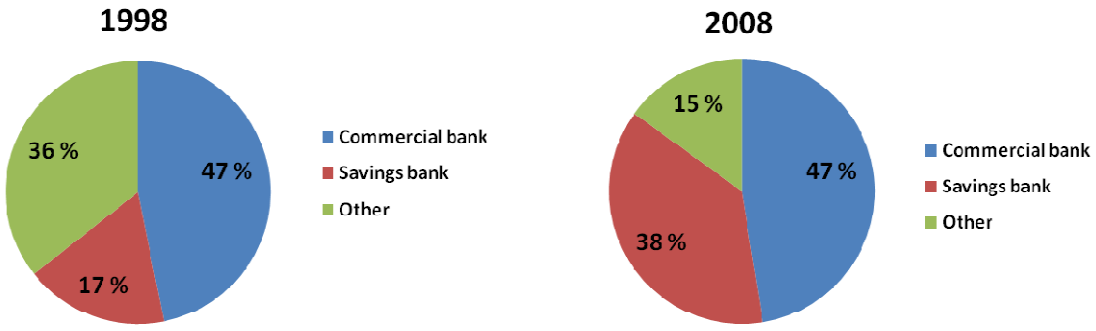


Figure 2: Distribution of relationships, by type of bank

This figure shows the distribution of relationships by bank category in 1998 and 2008. The bank category is defined by the Norwegian Ministry of Finance.

Note that the distribution in 2008 represents the relationships that have survived, and thus is ongoing in 2008[22]. In the appendix (Figure A) we include how the breakdown develops year by year during the sample period. It can be seen that the share of relationships with savings banks is increasing with time, the share related to commercial banks is stable, while the share related to *other* lending types is decreasing. One possible explanation for the

[21] In a similar study Ongena & Smith (1998b) find the same median duration of 4 years, and that 82% of the observations last for less than 8 years. In contrast to our study they had the opportunity to study the duration beyond 11 year. They found that less than 1% of the relationships lasted for more than 15 years. Note that their sample is limited to firm-bank relationships of firms listed on the Oslo Stock Exchange, and that their sample period was 16 years from 1979.

[22] These are the relationships that are right-censored. The impact of censoring will be discussed later in section 3.3, *Challenges with the data*.

decrease among the "other" category is that it is dominated by credit firms that in general are thought of as providers of short-term financing on a more random basis compared to the two other categories. To sum up, the figure suggests that relationships to savings banks in general are longer than the ones to commercial banks and other lenders.

Further we looked into whether a firm uses the same bank for deposits and borrowings. The results are presented in Figure 3.

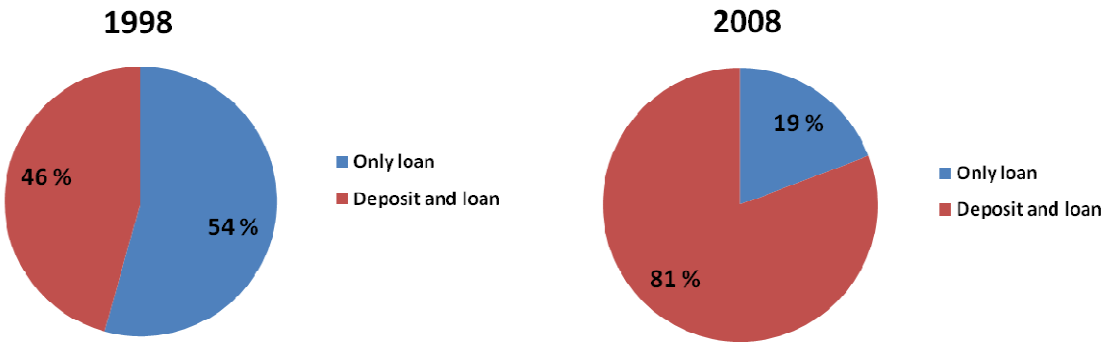


Figure 3: Distribution of relationships, by type of accounts held by the firm with the bank

This figure shows the distribution of relationships by the type of accounts the firm holds with the bank in 1998 and 2008. In the relationships categorized as "only loan", the firm did not have any deposits throughout the whole sample period. All numbers are obtained from the Dun & Bradstreet database.

In 1998, for 54% of the relationships, the firm holds both loan and deposit accounts with the bank, while in 46% of the cases the firm is only borrowing. Throughout the period the relationships where the bank provides both services seems to be strongest. The year-by-year development is presented in the appendix (Figure B). Further, on average in an "only loan"-relationship the amount of loan taken up is, 2.1 million NOK, while the average amount when the bank is providing both services is 2.2 million NOK.

As regards the ownership of banks, Figure 4 shows that, not surprisingly, most relationships in the sample are related to Norwegian owned banks. However, it is interesting to note that the share of internationally owned banks is increasing throughout the period, suggesting

these relationships to be stronger. More detailed information is provided in the appendix (Figure C).

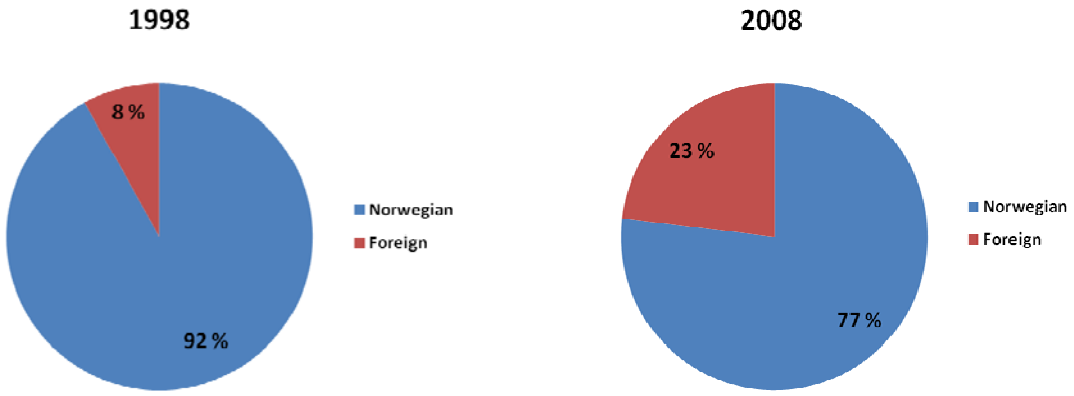


Figure 4: Distribution of relationships, by bank owner nationality

This figure shows the distribution of relationships by nationality of the bank owner. We have assumed that the nationality of a bank owner is the same during the period. All numbers are obtained from the Dun & Bradstreet database.

Finally in the overview Figure 5 shows the distribution of relationships by the age of the firm in 1998.

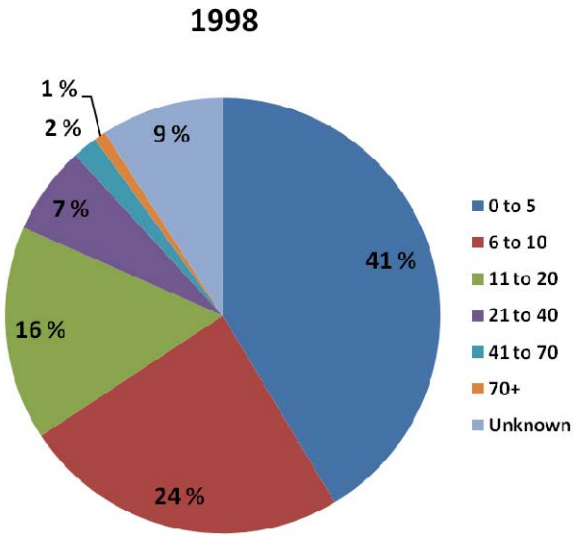


Figure 5: Distribution of relationships, by firm age

This figure shows the distribution of relationships by firm age in 1998. The age in 1998 is defined to be the number of years from the firm was founded (according to the registration in the Brønnøysund Register) till 1998. All numbers are obtained from the Dun & Bradstreet database.

We see that the majority (65%) of the relationships are the sample is related to firms that are younger than 11 years at the start of the sample, and the median age is six years. Among the still ongoing relationships in 2008 (the censored cases), the median age is 17 year. Note however, that since all firms naturally are ten years older, the median of 17 years is comparable to a median of seven years at the start of the sample period. This rather small one-year movement in the median may indicate that bank relationships are more durable for older firms. More detailed information can be found in the appendix (Table C).

In the appendix, Figures D to F show how the relationships are distributed with regard to geographical location of the firms and banks, together with the distribution of the industry in which the firms belong. Figure D shows that the majority of the relationships in the sample are related to firms operating in "Østviken" and in "Vestlandet". Figure E shows that the majority of the relationships are also related to banks located in "Østviken" and Vestlandet". This is representative for the overall distribution of where bank and firms are located in Norway[23]. Lastly, Figure F shows that the relationships in our sample are dominated by firms that operate within the service and commercial industries.

3.3 Challenges with the data

We will now discuss three issues that complicates the analysis of relationship duration.

The definition of a unique bank

In the sample, the firms maintain relationships with 285 different banks, and, like for the firms, we identify each bank by its unique organization number. These banks are either Norwegian owned banks or international banks with a branch in Norway. The number of unique banks is much larger than we first expected. According to The Norwegian Savings Banks Association there were 130 savings banks, 12 commercial banks and nine international banks operating in the Norwegian market in 2001.[24] This calls for a discussion around the definition of a unique bank in our sample. In Norway, banks are often divided into different legal entities. Hence, there exist multiple organization numbers for one unique bank. To give an example from our data, in 2007 the biggest lending bank by market share to

[23] Statistics Norway (Statistisk Sentralbyrå, SSB), 2008

[24] The Norwegian Savings Banks Association (Sparebankforeningen), 2001. *Nå kommer de utenlandske bankene.*

the corporate segment[25], the DnB NOR group, had eight different lending businesses, all with its own unique organization number.[26] Unfortunately, the available dataset did not enable us to adequately adjust for this, and this explains the high number of different banks in our sample. Yet, with our definition of a bank (identified by the organization number) we are able to report that the relationships in our sample are concentrated across a small number of banks. 55% of all the relationships at the beginning of the sample period are linked to a limited number of 15 banks. In 2008, 49% of the remaining relationships are linked to only five banks[27].

The fact that we incorrectly identify a bank as unique when it is actually just one of many lending branches in a bank has an impact on our duration analysis. For example if a firm switches its borrowing from one branch to another within the same bank, we define it as a termination. Possibly, we also wrongly identify a termination when a firm has several loans in different branches within the same bank. If the firm terminates its relationship with one of the branches in the real world, it can still have an ongoing relationship with the other branch(es). This shows that we potentially shorten the length of some relationships. However, this issue is heavily mitigated by the fact that almost 90% of the firms in our sample have concentrated all their loans to one bank (see Figure 1).

Mergers and Acquisitions

Between 1998 and 2008, nearly 30 mergers and acquisitions took place in the Norwegian bank market[28]. As mentioned before, the DnB Nor group is a dominant player in the Norwegian bank market, and was involved in four of the mergers and acquisitions. (See Figure G in the appendix). The largest merger took place between DnB and Gjensidige Nor in 2004. As regards the duration of relationships, it is important to discuss mergers and acquisitions. It is reasonable to argue that when customers of merged banks continue with the combined entity, the firm-bank relationship also continues. In our study this is only partly taken into account. In most cases only the organization number of one of the banks in

[25] The Norwegian Financial Services Association (Finansnæringens Hovedorganisasjon, FNH).

[26] DNB NOR: Bank, Cresco, Kort, Lindorf, Postbanken,Verdipapir, Kreditt, Finans.

[27] By looking up the organization numbers of these 15 banks on the web, it becomes clear that the firms in our sample are mainly concentrated around the banks that from 1998 to 2008 had the biggest lending market share to the corporate sector. (Market share 2008: DnB Nor Group 30.7% Nordea 9.9%, Fokus Bank 4.0%, Handelsbanken 2.7%) Source: The Brønnøysund Register/FNH.

[28] The Financial Supervisory Authority of Norway (Kredittilsynet), 2005.

the merger remains unchanged, while the other bank's number is deleted.[29] This means that all relationships that are related to a (bank) organization number that is deleted will be recognized as a terminated relationship in our study. Clearly, our analysis will be sensitive to this; the duration of some relationships is identified as shorter than they actually are. However, it is not clear how many of the relationships in our sample are affected by mergers. Still, since we know that the majority of the relationships are limited to a handful of different banks, it is reasonable to assume that mergers and acquisitions in the bank market have the same impact on all types of firms.

What empirical implications are the mergers and acquisitions likely to have for our study? First we assume, according to our search at the Brønnøysund Register, that the organization number of the (smaller) acquired bank, or the smallest player in a merger, is deleted. Further we assume that savings banks tend to have more small and medium sized firms in their client portfolio, compared to the bigger commercial banks[30]. Then it is likely that there is a bias in our study towards that the duration of relationships of smaller firms are being identified as shorter than they actually are. On the other hand, in a survey from 2003 done by Bedriftsforbundet, bigger players such as DnB, Gjensidige and Nordea had 18%, 14% and 13% of the market share in the small-medium sized corporate market, respectively[31]. The relatively large market shares among these commercial banks should mitigate the mentioned bias.

Case study on Mergers and Acquisitions

We further examine the sensitivity of the duration of bank relationships through a case study of those relationships in our sample that are related to Gjensidige.

As already mentioned, the largest merger that took place during our sample period was the one between DnB and Gjensidige Sparebank Nor in 2004.[32] In this case the organization

[29] By making a few organization number checks at the Brønnøysund Register (www.brreg.no) we noticed that it is the organization number of the acquired bank that that tends to be deleted. In mergers it seems like the biggest bank's number continues to be used.

[30] This is supported by a study done by Bedriftsforbundet in 2003, where the majority of small and medium sized firms preferred to have a smaller regional or local bank as their lending bank.

[31] The survey was done before the merger between DnB and Gjensdige Sparebank Nor (The Norwegian Savings Banks Association, 2004).

[32] Dnb Nor (www.dnb.com).

number of Gjensidige Sparebank Nor was deleted.[33] In our sample, 709 relationships are related to this number in 1998, and, in 2003, 449 of these relationships are still maintained. In our study these relationships are recognized as terminated in 2004 (since the organization number no longer existed by then). Compared to the general distribution of observed durations of bank relationships (Table 2), it is very unlikely that all of these relationships were terminated in 2004 due to the merger. Actually, by using the new organization number, we manually looked up and found that 403 of the relationships were maintained in 2004, which exemplifies the sensitivity of duration of relationships to mergers in our study.

Recall, however, that we argued that mergers will affect all kinds of relationships (regardless of size, age etc). Since we haven't empirically analyzed the impact of a merger in our study we refer to a study by Sapienza (1998). According to her, a bank merger can weaken the benefits from a bank relationship, culminating in the end of the relationship. In her study she finds that a firm is more likely to end a relationship after a merger. Even though Sapienza's study did not control for relationship duration, her study suggests that some of the problems in our study regarding mergers, are reduced by the fact that mergers tend to have a negative effect on the value of the relationship. One probable reason for a reduced relationship value is that a merger or acquisition may cancel out the private information of banks due to new processes of providing loans, and change of staff.

Right-censoring

Since we are not able to observe the outcome of relationships that continue as of 2008, our sample is affected by right-censoring. This is the case for 16% of the observations. We recognize that (also) this can bias inferences about the duration of relationships.

Recall, that if (for each of the Figures 2-4) the share of one characteristic is increasing over the period from 1998 to 2008, this suggests that this characteristic has a positive impact on the duration of the bank relationship. However, the presence of right-censoring leads to inexact comparisons among characteristics. For example, Figure 4 shows that the share of relationships with internationally owned banks is increasing greatly from 1998 to 2008, suggesting that internationally owned banks lead to longer relationships. However, it is impossible to measure the exact effect (compared to Norwegian owned banks) as long as we

[33]The Brønnøysund Register.

cannot observe the further development of the censored cases. Hence, there is a chance that the positive impact of censored characteristics is underestimated. The table below shows a list of characteristics suggested to have a positive impact on maintaining the relationship. Due to right-censoring there is a chance that the positive impact for these characteristics is underestimated.

Table 5: Characteristics of which the positive impact on relationship duration may be underestimated

Category	Characteristic
Bank	"Savings Bank"
Account type	"Deposit and Loan"
Bank ownership	"Foreign"

This table lists, for each category the characteristics that has an increasing share of relationships in 2008 compared to in 1998, suggesting that these characteristic have a positive impact on relationship duration. The positive impact of these characteristics may be underestimated due to the presence of right-censoring in our data sample.

The extent of this issue depends on whether the censored observations are concentrated around certain firm, bank and/or relationship specific characteristics. For example, if all the censored relationships are related to savings banks, or if all the censored cases are associated with old firms, this will lead to a more severe bias than if the censored cases are more equally distributed among different characteristics. To examine this we look back at the Figures 2-4 which for the 2008 part shows the distribution of the censored relationships by the different characteristics[34]. The characteristics represented in 1998 are in general represented with a similar share in 2008, suggesting that the censoring bias is moderate.

3.4 Determinants of relationship duration

Motivated by the theoretical literature on bank relationships, we incorporate firm-, bank- and relationship specific data that, in addition to time itself may influence the duration of-, and likelihood of ending a bank relationship.

[34] In particular we observe that compared to “other” bank categories, savings- and commercial banks, are more affected by censoring. Further, “Only loan” is the most censored account type, and Norwegian owned banks represent the biggest share of censored cases regarding ownership.

Ongena & Smith (1998b) find in a similar study that small, young and highly leveraged firms and firms with multiple bank relationships tend to maintain the shortest relationships. To do a comparison to their study, we first build a model with only explanatory variables that serves as a proxy for these characteristics. In addition, based on theory and our own hypotheses/our own critical review, we build three models where we add other firm-, bank- and relationship specific indicators[35].

Size and Age

Size and age are important factors when analyzing a bank relationship. Fama (1985), Diamond (1984, 1991) and Rajan (1992) found that small and young firms are most dependent on bank financing due to asymmetric information. Further, Petersen & Rajan (1994) found that older firms tend to receive the best conditions on their loans. *Ln Income*, defined to be the natural logarithm of year-end total income of a firm, serves as a proxy for firm size. To measure the firm's age, we define *Age* to be the number of years since a firm's founding date (as registered in the Brønnøysund Register). However, this measure causes problems for our duration model i.e. the spurious correlation that firms become older as duration lengthens. To cope with this in our regression model, we therefore introduce the variable *Age At Start*, defined to be the number of years from a firm was founded till the start of our sample (1998).

Profitability

According to a study by Titman and Wessels (1988), highly profitable firms are less dependent on bank loans and other external financing, which is consistent with the Pecking Order theory. It is also the most profitable firms that in general are offered the best lending rate (Rajan, 1992). We construct *Profitability* by using the ratio of (gross) operating result[36] to the year-end book value of assets.

Leverage

Firms in Norway are heavily reliant on bank debt[37], and a highly leveraged firm is more dependent on bank financing than a mostly equity-financed firm, possibly making it costlier for the firm to switch banks. However, Farinha & Santos (2000) found that (small) highly

[35] An overview of the different models is provided in the result section.
[36] Sales minus cost of goods sold.
[37] Statistical Yearbook of Norway 2008(Statistics Norway, SSB).

leveraged firms are more likely to initiate multiple relationships, and with that to switch a relationship faster than others. We generate *Leverage* by using the firm's book value of debt divided by the sum of year-end book value of total assets given in the Dun & Bradstreet database.

Multiple bank relationships

Various studies show that firms choose to initiate multiple bank relationships, even though it is not necessarily purely advantageous for the firm.[38] However, firms with multiple bank relationships have more than one potential source of inside bank financing and should therefore be less exposed to holdup threats, thus finding the cost of ending one bank relationship to be lower. Also, competition between the banks may cause the firm to shift all services to one bank and end services with another (Ongena and Smith, 1998b). The dummy variable *Multiple Relationships* equals one when a firm maintains more than one potential source of inside bank relationship, and zero otherwise.

Tangibility

A way to determine the riskiness of a firm is by looking at its ratio of tangible assets. Tangible assets can be seen as firm security or collateral. Firms with fewer tangible assets are more likely to experience greater information asymmetry when communicating their value to outside investors and therefore a greater degree of financial constraint (Bhagat et al., 2005). We measure *Tangibility* by dividing a firm's fixed assets by the year-end total book value of assets.

Creditor Concentration

Petersen & Rajan (1994) showed that firms have a tendency to concentrate their borrowings, even when a firm has multiple lending sources. We have included the Herfindahl-Hirschman index (HHI) as a measure of the *Creditor Concentration*. This is calculated by summing the squared loan shares of each firm per year[39]. The index produces a ratio between 0 and 1. High HHI index indicates high creditor concentration, meaning that the firm has a large share of its loans in one single or just a few banks. Our hypothesis is that firms with higher credit concentration can obtain better loan conditions, which implies longer relationships.

[38] See Houston & James (1996), Guiso & Minetti (2004) and Rajan (1992).

[39] $HHI_{creditor\ concentration} = \sum_{i=1}^{N} s_i^2$, where s_i equals loan divided by total loans of firm *i* (Ongena et al., 2007).

Deposit & Loan

Deposit & Loan equals one when a firm holds both deposit and loan accounts with the bank and zero if the firm has only a lending relationship with the bank. Our hypothesis is that more services provided/used also lead to less information problems, and thus longer duration of relationships.

Bank Category

In general, Kim et al. (2004) find that firms prefer to take up loans in large and diversified banks. However, the Bedriftsforbundet's study from 2003 on the Norwegian market finds that the majority of small and medium sized firms prefer smaller regional or local banks as lenders[40]. In our model *Bank Category* divide the banks in our sample into three categories, savings banks, commercial banks and other lenders[41].

Nationality

From the descriptive analysis we found that the relationships in our sample can be divided into two groups dependent on whether the ownership of the bank is Norwegian or not. To analyze whether the nationality of the bank owner has an impact on the length of relationships, we create a dummy *Nationality* that equals one when the bank is Norwegian owned and zero otherwise. Based on our intuition it can be argued that relationships to Norwegian owned banks are expected to be the strongest due to the fact that Norwegian banks traditionally have provided the largest share of bank financing to Norwegian firms. At the same time, as we discussed earlier, the competition from international banks in the Norwegian lending market has increased during the sample period. Due to lower brand value in Norway, we can assume that international banks must offer better loan conditions to gain market share. Thus, such better conditions can serve as a counter-argument for expecting Norwegian owned bank relationships to be the strongest.

The table below contains summery statistics on the bank relationship characteristics previously described. Recall that we obtained yearly values from the Ministry of Finance and the Dun & Bradstreet database[42]. Note however, that in the table below we have used the

[40] The Norwegian Savings Banks Association (Sparebankforeningen), 2004.
[41] The category "Other lenders" is dominated by public banks, insurance firms and credit firms. The category division is correspondent with the one in the Dun & Bradstreet database.
[42] Strictly speaking, we generated *Ln Income, Age At Start, Profitability, Multiple Relationships, Tangibility, Creditor Concentration, Primary Bank* and *Deposit & Loan*.

mean value of each relationship characteristic throughout the sample time.[43] To give an example: if a firm had a relationship to a bank for three years, with leverage equal to 0.5, 0.6 and 0.45 in 1998, 1999 and 2000 respectively, the leverage measure for this observation is 0.52 (the *mean* of 0.5, 0.6 and 0.45).

To adjust some of the extreme values which are present in the data set and make the sample more robust to outliers, we use the Winsorization method. By this, we apply a 98% Winsorization. This means that observations above/below the 99th/1st percentile of each tail are set to the 99th/1st percentile.[44] The implication is that the sample becomes more evenly distributed.

Table 6: Descriptive statistics for explanatory variables

Variable	N	Mean	Median	St.Dev	Minimum	Maximum
Firm specific						
Income	8,057	22,316	3,891	71,373	0	548,359
Ln Income	7,893	8.2	8.2	1.8	0	13.2
Age	8,613	11.7	8	12.9	0	72.5
Age At Start	8,613	9.6	6	11.6	0	70
Profitability	8,054	-0.02	0.05	0.43	-3.06	0.65
Leverage	8,054	1.1	0.8	1.4	0	12.6
Multiple Relationships	10,130	0.1	0	0.3	0	1
Tangibility	8,053	0.4	0.4	0.3	0	8.0
Creditor Concentration	9,860	0.98	1	0.08	0.29	1
Bank specific						
Nationality	10,130	0.87	1	0.34	0	1
Relationship specific						
Deposit & Loan	10,130	0.46	0	0.49	0	1

For this table we use the mean value of each variable throughout the sample period. All values are obtained or derived from the Ministry of Finance or the Dun & Bradstreet database. Income is the total income of the firm in 1,000 NOK. Ln Income is the natural logarithm of yearly total income. Age is years between the firm's founding date and the time of measurement. Age At Start is the years between a firm's founding date and the beginning of the relationship (1998). Profitability is the ratio of operating income to the year-end book value of assets. Leverage is the book value of debt divided by year-end value of total assets. Multiple relations equal one when a firm maintains more than one bank relationship and zero otherwise. Tangibility is fixed assets divided by the year-end total book value of assets. Creditor Concentration is measured by the HHI index and is calculated for each firm by summing the square of loan shares. Finally, Deposit & Loan equal one when a firm has both a deposit and loan account in the bank and zero if the firm has only a lending relationship with the bank. N is the number of relationship observations. Note that for some of the variables, due to missing values in the Dun & Bradstreet database, the number of observations is lower than the total sample size of 10,130.

[43] This is consistent with the methodical/econometrical specifications defined in section 4.

[44] An important note is that there are some weaknesses in the accounting database (i.e. wrong/ illogical values, negative or zero assets/D+E etc) which would have affected the further analysis by inducing some extreme/remarkable results. This is especially evident for Income Age At Start, Profitability and Leverage which are the variables we Winsorize. In the appendix Figures H-J show histograms of the sample for Profitability, Leverage and Age. The histograms show the distribution of the sample before and after implementing the Winsorizing method, confirming that the method adjusts for extreme values.

The average firm in our sample generates total income of 22 million Norwegian Kroner, while the median of firm size is 3.9 million NOK. As we have not imposed any restrictions on the size of the companies of our initial sample, our selection contains both very small and very large companies. Ten percent of our selected sample have total income of less than 500,000 NOK. A few large firms are included in the sample, and before adjusting for Winsorization the five percent largest firms have an average total income of 850 million NOK.[45]

The average firm age in 1998 is 9.6 years, while the median is only six years. The low average firm age indicates that most of the firms are young or recently established firms. Before Winsorization, the ten percent oldest firms averaged 38 years, but Age-at-start (which is Winsorized) shows that the oldest firm of the adjusted sample is 70 years in 1998.

The average firm has a negative profitability of 1.7% and a debt to asset ratio of 1.05. 30% of our sample have negative profits, while the remaining 70% generate an average positive gross profit of 12.6%. A negative gross profit means that a firm uses more money to produce their goods or services than they can collect in sales.[46] However, as shown in the above table, the median profitability is 5%. A debt ratio greater than one means that a firm has negative equity value. Negative equity implies that the value of a firm's debt is bigger than the value of its assets. 35% of the total sample have a debt ratio greater than one. Based on theory and empirical studies, we know that young, small firms are usually highly leveraged.

Firms which maintain multiple bank relationships during the sample period generate 11% of the firm-bank observations. We count each firm-bank relationship as one observation; a firm with multiple relationships enters as two or more observations. We do not account for late entries, which means that the firms registered with multiple relationships did in fact establish all their relationships in 1998.

Our sample has an average tangibility (liquidity) ratio of 0.45. A low liquidity ratio means that a firm has a small amount of collateral in terms of assets. The firm is likely to experience information asymmetry when signaling their firm value.

[45] As mentioned previously, we scale the size by taking the natural logarithm of income.
[46] Although not shown here, the average gross profit of a sample firm seems to be increasing through time; the average profit was 5.5% in 1998, while this increased to 12.6% in 2008, but these numbers are survival biased.

An average creditor concentration of 0.98 is very high. This means that most firms have concentrated their relationship around one lending source. This supports the previous finding in the descriptive section, that most Norwegian firms maintain only one bank relationship.

Of the relationships 87% are with a Norwegian owned bank and 45% of the relationships are a lending relationship only.

Table 7: Correlation matrix

Variable	Ln Income	Age At Start	Profitability	Leverage	Multiple Rel.	Tangibility	Creditor Conc.	Nationality
Age At Start	0.16***							
Profitability	0.21***	0.09***						
Leverage	-0.18***	-0.09***	-0.57***					
Multiple Rel.	0.11***	-0.04***	-0.04***	-0.01				
Tangibility	-0.21***	0.05***	0.05***	-0.07***	-0.001			
Creditor Conc.	-0.04***	0.05***	0.01	0.02	-0.64***	-0.03**		
Nationality	-0.09***	-0.05***	-0.04***	0.02*	0.02**	0.02**	0.02*	
Deposit & Loan	-0.04***	-0.03***	-0.06***	0.06***	-0.08***	-0.07***	0.11***	0.01

*Significant at 10%, **significant at 5%, ***significant at 1%

This table shows the pairwise correlation matrix where a bivariate t-test determines the statistical significance of the coefficients.

The correlation matrix describes the degree of the relationship between two variables of our model. It is evident that several of the variables are correlated. For instance; profit and leverage have a negative correlation coefficient of 0.57, which is very high. However, this is economically reasonable as the leverage of a firm tends to decrease as profit increases.

Given the method for constructing the variables, it is not surprising that some of the variables are highly linearly correlated. To what extent this correlation influences the regression estimates remains an empirical issue.

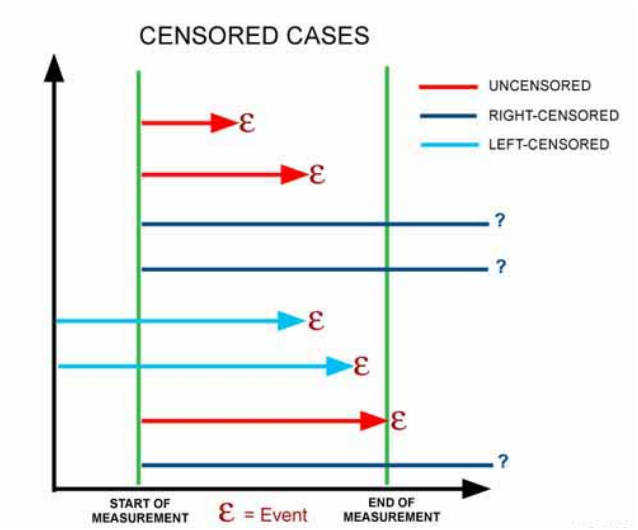
4. Methodology and Econometric specification

4.1 Survival and duration analysis

Survival and durational analysis is concerned with studying the time that passes before the occurrence of an event. Originally the survival analysis was concerned with treatment of patients until death, which accounts for the name given to these methods. However, the analysis is applicable to many areas, and in our case we are using the method to analyze the duration of firm-bank relationships.

Survival time and censoring

The survival time, *T*, may be thought of as a non-negative random variable. *T* represents the duration of time that passes before the occurrence of a certain random (failure) event.[47] Censoring is a special characteristic of survival data, and it occurs when we are not able to observe the true duration of the survival time. Failure to take censoring into account can produce serious biases in estimates of the distribution of survival time and related quantities. The figure below shows different censored cases.



Source: Garson (2009)

Figure 6: Censoring

[47] Cleves, M.A., W.W. Gould and R.G. Gutierrez. (2004). *An Introduction to Survival Analysis Using STATA(Ch.1 p. 1-3)*. Stata Corporation, Texas.

Right-censoring is defined as the point in time when the failure event occurs and the subject is no longer under observation. Recall that our sample period is from 1998 to 2008. However, some relationships are still ongoing at the end of 2008, and those observations are right-censored cases. Given the procedure of our study –it covers a limited time period - the censoring is fixed (as opposed to random).

Further, *left-censoring* applies when we do not know for how long an observation has been at risk before the start of the measurement. Recall our discussion in the section about sample selections. We have available relationship data from 1997, but some of the relationships had probably already lasted for some years at that time (1997). In other words it is not known when these relationships started, thus they are left-censored cases. Controlling for left-censoring is challenging, and thus, according to Ongena & Smith (1998b), it is often ignored in duration analysis. Nevertheless, Heckman and Singer (1984) argue that the biases induced by left-censoring can be as severe as those created by right-censoring.

For those familiar with survival analysis we would like to mention that there is no presence of interval censoring or truncation in our study sample. Given the absence of these characteristics in our case, we do not pursue this topic[48].

Survivor function and Kaplan and Meier estimator

There are several ways of presenting the behavior of *T*, but it is common to describe *T's* survivor function *S (t)*:[49]

$$S(t) = Pr(T > t) \qquad (1)$$

The survivor function reports the probability that the survival time lasts beyond time *t*. The function is equal to 1 at *t=0* and decreases toward zero as it goes to infinity.

[48] For a more general understanding of interval censoring and truncation we recommend *An introduction to survival analysis* (2004) by Cleves, M.A., W.W. Gould and R.G. Gutierrez. 2004 (Ch.4. p.32-36).

[49] Which is nothing more than the reverse of the cumulative distribution function *F(t)=Pr(T≤t).* Thus, *F(t)=1-S(t).*

The estimator of Kaplan and Meier (1958) is a nonparametric common method of estimating a preliminary sketch of the survival function:

$$\hat{S}(t) = \prod_{j|t_j \leq t} \left(\frac{n_j - d_j}{n_j} \right) \qquad (2)$$

where $\hat{S}$(t) is the estimated probability that a subject will survive beyond time *t*. The n_j represents the number of subjects *at risk* at time t_j (the number at risk includes those for whom the event has not yet occurred) and d_j is the number of failures at time t_j. The product is less than or equal to *t* over all observed failure times. The estimator can estimate survival function even in the presence of right-censored cases. The number of failures d_j at time t_j will be the same with or without a correction for right-censoring, while the number of subjects at risk varies. Without a correction n_j is the number of subjects that are still remaining in the data set before time t_j. With a correction n_j becomes the total number of surviving subjects.

It is possible to construct approximate standard errors around the Kaplan-Meier estimate. The one we report is given by Greenwoods's (1926) formula:

$$\widehat{Var}(\hat{S}(t)) = \hat{S}^2(t) \sum_{j|t_j \leq t} \frac{d_j}{n_j(n_j - d_j)} \qquad (3)$$

It is important to be aware that the Kaplan and Meier estimator is preliminary in the sense that it is assumed that the event probabilities depend only on time. In other words, the estimator does not take into account any covariate effects. Further it is assumed that all subjects behave similarly and the computed survivor function is assumed to describe all subjects, also the censored cases. Even given these rather strong assumptions, the Kaplan and Meier estimator can still be fruitful in the exploratory stages of our study.

The Proportional Hazard function

Another commonly used function within survival analysis is the hazard function *h(t):*

$$h(t) = lim_{\Delta \to 0} \frac{Pr\,(t+\Delta t > T > t | T > t)}{\Delta t} = \frac{f(t)}{S(t)} \qquad (4)$$

where *f(t)* is *T's* probability density function[50]. The hazard function is the instantaneous rate of failure. As can be seen from equation (4) the (limiting) probability is that the failure event occurs in a given interval, conditional upon the subject having "survived" to the beginning of that interval, divided by the width of the interval. The hazard function can vary from zero to infinity (zero meaning no risk).

Even though neither the survivor function nor the hazard function provides additional information that could not be obtained directly from the density function *(f)*, in economic terms it is often more convenient to talk in terms of these functions.

Over time the hazard function can increase, decrease, remain constant, or even take more serpentine shapes. When *h(t)* is increasing over time, the function is said to have positive duration dependence. In the same way, when *h(t)* is decreasing over time, negative duration dependence occurs. Finally, when there is no relation between *h(t*) and *t*[51], we are talking about constant duration dependence. This means that the likelihood of failure is the same throughout time.

There are several hazard models, but the so-called proportional hazard model specification is particularly popular because it is econometrically convenient to use:

$$h_j(t) = h_0(t)exp(\beta_0 + x_j\beta_x) \quad (5)$$

As the equation shows, with this specification the hazard function is composed of two separate parts, multiplied. The first part, h_0, is the baseline hazard which is exclusively a function of duration time. In other words, the baseline determines the shape of the hazard function with respect to time. The second part is a function of explanatory variables other than time. It is traditionally chosen to take the form $exp(\beta_0 + x_j\beta_x)$, where *x* is a vector of observations on the characteristics of a subject at risk, *j*, (which may vary with time) and β is a parameter vector. The function *exp ()* was simply chosen to avoid the problem of $h_j()$ ever turning negative. In addition, the model offers an appealing interpretation of β since the logarithm of $h_0(t)$ is linear in *x*, which means that β reflects the partial impact of each variable in *x* on the log of the estimated hazard rate.

[50] *f(t)=-S'(t)*

[51] $\frac{dh(t)}{dt}=0$

The key issue is that time itself is separated from the explanatory variables so that the hazard function is obtained simply by shifting the baseline hazard as the explanatory variables change, i.e. for all subjects the hazard function is proportional to the baseline hazard function (Kennedy, 2003)

Semiparametric modeling

In our study we exploit the fact that within survival analysis there are methods that do not require assumptions about the distributions of failure times. These methods are referred to as semiparametric models[52]. With survival data, the key insight into removing the distributional assumption is that, because events occur at given times, these events may be ordered and the analysis may be performed using the ordering of the survival times exclusively.

The Cox proportional hazards model

Recall that the nonparametric estimator Kaplan-Meier is based only on time dependence. However, the most advanced part of our study examines the relationship between survival – in the form of the hazard function – and a set of explanatory variables (or covariates). The Cox proportional hazards regression model (introduced by David Cox in 1972) does take covariates into account. It assumes that the covariates multiplicatively shift the baseline hazard function:

$$h_j(t) = h_0(t)exp(x_{j1}\beta_1 + x_{j2}\beta_2 + \ldots + x_{jk}\beta_k) \qquad (6)$$

where $x_{j1} + x_{j2} + x_{jk}$ are the values of the covariates for the *i*th subject, and the brackets represent the linear predictor. Notice that there is no constant term (intercept) in the linear predictor: The constant is absorbed in the baseline hazard.

Further, the regression model asserts that the hazard rate for the *j*th subject in the data is:

$$h(t|x_j) = h_0(t)exp(x_j\beta_x) \qquad (7)$$

[52] The models go under the name semiparametric because, as far as time is concerned, they are nonparametric, but since we are still parameterizing the effect of explanatory variables, there exists a parametric component to the model. (Cleves, Gould and Gutierrez, 2004, Ch. 9).

where the regression coefficients β_x are to be estimated from the dataset. Note that the Cox model does not use maximum likelihood estimation, but rather a maximum partial likelihood method that requires only the order of survival times to be known (when estimating the hazard ratio). Actual survival times are not used in partial likelihood estimation of the hazard function. This is why the model bases estimation of β_x on the ordering of the survival times[53].

In the model, the baseline hazard h_0 is not only estimated without any functional form, it is actually left unestimated. According to, Cleves, Gould and Gutierrez (2004), this is one of the reasons why the model is the most used model within survival analysis. The fact that the model makes no assumptions about the shape of the hazard function over time increases the computational feasibility. A key assumption of the model is, however, that whatever the shape, it is the same for all subjects. [54] In other words, the assumption is that hazards (or failures) are proportional: the hazard ratio will remain constant over time. Note that proportional hazards mean that hazards (failure events) are proportional over time, *not* that they are the same over time. (Note also that hazard rates are not hazard ratios, and their respective interpretations differ. See appendix section 8.2 for hazard definitions.)

Parametric modeling

Parametric models use probabilities that depict what occurs over the whole time period for each subject, given what is known about the subject during this time.

Recall that in the equation for the proportional hazard model (4) the functional form *exp()* was chosen. In general it could be any non-negative function, Φ:

$$h_j(t) = h_0(t)\Phi(\beta_0 + x_j\beta_x) \qquad (8)$$

What is important to know about parametric proportional hazard models, is that these models require specification of function form of the baseline hazard, h_0 and Φ[55]. Thus, in

[53] In our study we identify failures once a year, meaning that each year there are tied failures. Ideally, partial likelihood methods would have no tied data, but a rather simple ordering of failure times. To handle this we use partial likelihood algorithms that have been adapted to handle ties. These are available in statistical software (e.g. STATA).

[54] One subject's hazard is a multiplicative replica of another's (Cleves, Gould and Gutierrez,2004, Ch. 9).

[55] Recall that in the semiparametric Cox model the baseline hazard does not need to be specified.

parametric models, the shape of the distribution of survival times arises from the specification of the baseline hazard.

The exponential model

The exponential model is commonly used. It is the simplest parametric model because it assumes that the baseline hazard is constant,

$$h(t/x_j) = h_0(t)exp(x_j\beta_x) \qquad (9)$$

$$= exp(\beta_0)\, exp(x_j\beta_x)$$

$$= exp(\beta_0 + x_j\beta_x)$$

for some constant β_0. In the model, $h_0(t)$ being constant means that the failure rate is independent of time. Also, since the model is proportional, and the baseline hazard is a constant, whenever the risk is doubled or tripled, the new risk is still constant, just higher (Rodríguez, 2009).

The Weibull model

Another much-used model is the Weibull model. The model assumes a baseline hazard of the form:

$$h_0(t) = pt^{p-1}\, exp(\beta_0) \qquad (10)$$

for $p>0$, where p is some ancillary shape parameter estimated from the data and the scale parameter is parameterized as $exp(\beta_0)$. Given a set of explanatory variables x_j, under the proportional hazards model,

$$h(t/x_j) = h_0(t)exp(x_j\beta_x) \qquad (11)$$

$$= pt^{p-1}\, exp(\beta_0)\, exp(x_j\beta_x)$$

$$= pt^{p-1}\, exp(\beta_0 + x_j\beta_x)$$

The shape of the hazard function is determined by the estimated parameter p. Note that the Weibull distribution allows for duration dependence. When $p = 1$, this model reduces to the exponential and has constant risk (hazard) over time. If $p > 1$, then the risk (hazard) increases over time. If $p < 1$, then the risk (hazard) decreases over time.

If we pick the Weibull distribution as a baseline risk and then multiply the hazard by a constant, γ, in a proportional hazards framework, the resulting distribution turns out to still be a Weibull (Rodríguez, 2009).

4.2 Comparison of the survival models

To examine the influence of duration and explanatory variables on the likelihood of ending a relationship, the starting point for our study has been the Cox model. According to Fox (2006) the fully parametric hazard models have been superseded by the Cox model.

It is challenging to make reasonable assumptions about the shape of the hazard functions. Compared to the parametric models, the advantage of the Cox model is that we do not need to make assumptions about the baseline hazard $h_0(t)$, assumptions about which we might be wrong and which, if we were wrong, could produce misleading results about the regression coefficients, β_x. The cost is the same as for other statistical tests i.e. if more information was available, in this case if we knew the functional form of $h_0(t)$, we would be able to do a better job. Actually, Fox argues that not having to assume a possibly incorrect form for the baseline hazard more than makes up for small inefficiencies in estimation.

Further, the information we are trying to obtain must also be taken into consideration when choosing between a semiparametric and a parametric analysis. Cleves, Gould and Gutierrez (2004, p. 232) argue that if the focus of the study is the parameter effects (hazard ratios), rather than the prediction of time to failure, it is probably favorable to use the semiparametric approach. In our study we use maximum likelihood methods for the three models (Cox, exponential and Weibull) and compare the results.

4.3 Assumptions about the explanatory variables

According to Kabelfleisch and Prentice (1980) the variables are required to be either *defined* or *ancillary* with respect to survival time. The age of a firm is an example of a defined variable, since it follows a deterministic path, meaning that its path varies deterministically with the length of the bank relationship. This path is known before the firm enters the relationship. On the other hand, an ancillary variable is independent from the duration of survival time, and follows a stochastic path.

We previously discussed in chapter 4.1 the fact that the explanatory variables in our model may vary over time. In order to take this into account, for each sample observation in our study (each firm-bank relationship) we use the mean value of each of the explanatory variables (relationship characteristics) throughout the survival time. By building our model in this way, we assume that the conditional likelihood of ending a relationship depends on the historical values of the variables.[56]

[56] We also built models where we assumed conditional likelihood only dependent on the value of the variables at the time of failure. These models gave similar results as the one reported, but lead to more missing values since the richness of accounting data varies from year to year. For all models, category variables such as *Nationality* and *Bank Category* we have copied the information from year to year when data are missing. By doing this we assumed category variables to be "non-varying", i.e. to be the same during the whole sample period.

5. Empirical results

We start by showing the raw cumulative probability of a relationship length. Figure 7 shows a smoothed estimate of the survivor function, $\hat{S}(t) = Pr(T > t)$, using the Kaplan and Meier estimator which is described in detail earlier in the methodology section.[57]

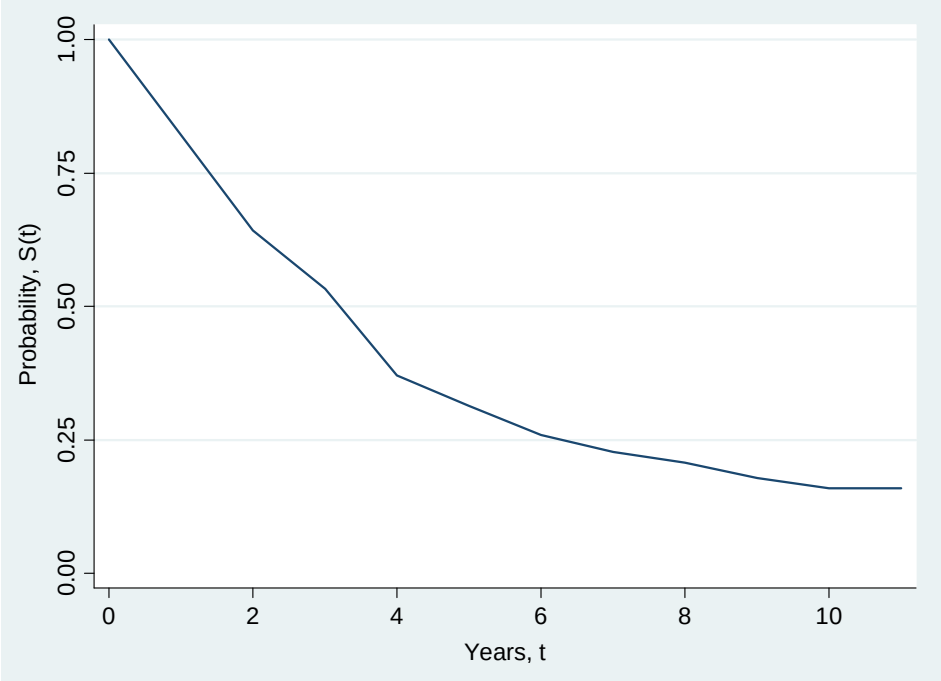


Figure 7: Kaplan-Meier survival estimate (smoothed)

This figure shows the non-parametrically estimated survivor function. The function is constructed by using the Kaplan and Meier (1958) estimator, $\hat{S}(t) = \prod_{j|t_j \leq t}\left(\frac{n_j - d_j}{n_j}\right)$ *where Ŝ(t) is the estimated probability that a relationship survives beyond year t.*

As stated in section 4, the Kaplan-Meier depends only on time, and the estimator does not account for covariate effects. In our case, the survivor function decreases quickly the first four years. The estimated likelihood of a relationship surviving beyond the first year is 82%, while the estimated chance of surviving past four years is only 37%. This means that 18% of the relationships started in 1998 fail already during the first year and another 18% fail the

[57] $\hat{S}(t)$ is the estimated probability that a relationship survives beyond time *t* and d_j is the number of relationships leaving the data set during year *j*. n_j is the number of relationships that are still remaining in the data set before time t_j.

second year. As we can see from the figure the survivor functions tend to be decreasing over time (for more detailed information, see Table D in the appendix. The appendix also provides an overview of estimated survivor functions for subgroups in Figures K to N). This suggests that the hazard function is increasing, causing positive duration dependence. The presence of positive duration dependence will be more formally discussed in the next section when we introduce explanatory variables to our analysis.

5.1 Cox partial likelihood estimates

We estimate the impact of the firm-, bank- and relationship-specific characteristics on the conditional probability of a bank relationship ending using the proportional hazard specification described in Equation (5) and the semiparametric Cox (1972) partial likelihood model. In the regression tables we report the hazard ratios for each explanatory variable. The hazard ratio displays the relative risk or the probability of an event occurring in the exposed group versus the non-exposed group. This makes it intuitively easy to see the effect of a given variable on the relative hazard; a hazard ratio *below* one indicates that the hazard of the exposed group is less than that of the non-exposed group. Equivalently, a hazard ratio *above* one indicates a higher hazard relative to the non-exposed group. For instance; in *Model 1* (Table 9) the multiple relationships hazard explains multiple-bank firms to have 56% higher hazard ratio than single bank relationships, while in *Model 4* the corresponding hazard ratio implies multiple-bank firms to have 21% higher hazard than single bank firms. Recall from section 4 (Equation 7) that the connection between the hazard rate and the coefficient of a variable is as follows; $h(t|x_j) = h_0(t) * e^{x_j \beta_x}$ where $h_0(t)$ is the baseline hazard and β_x is the estimated coefficient.[58]

We present four different models in Table 8 below. The models are constructed from variables detailed in the descriptive analysis of explanatory variables.

[58] As the baseline hazard is typically "integrated out", or heuristically removed from consideration, calculating the coefficients from the hazards can therefore be explained by: $\beta_x = \ln h(t|x_j)$ when all other variables are zero.

Table 8: Overview of explanatory variables included in our survival analysis

	Model 1	Model 2	Model 3	Model 4
Base firm specifics	Size (LnIncome) Age At Start Profitability Leverage Multiple Relationships	Size (LnIncome) Age At Start Profitability Leverage Multiple Relationships	Size (LnIncome) Age At Start Profitability Leverage Multiple Relationships	Size (LnIncome) Age At Start Profitability Leverage Multiple Relationships
Extended firm specifics		Tangibility Creditor concentration	Tangibility Creditor concentration	Tangibility Creditor concentration
Bank specifics			Bank nationality Bank category	Bank nationality Bank category
Relationship specifics				Relationship type

This table provides an overview of the different models used to build the final survival model, Model 4. All the explanatory variables are described in section 3.4.

Model 1 is based on Ongena & Smith's study (1998b) and other previous literature. The model includes proxies for the following explanatory variables: size, age, profit, leverage and multiple relationships. This model will serve as a foundation for the further regressions and testing of additional variables. *Model 2* includes additional firm-specific variables based on our own critical assessment of what may impact the relationship duration, and is motivated by the descriptive statistics section. The additional variables are creditor concentration and tangibility (or liquidity). *Model 3* upholds all the firm-specific variables, and now includes bank-specific variables. These are nationality of the bank (domestic or foreign) and bank type (commercial, savings or other). As can be seen in the table above, *Model 4* still originates in the extended base model and bank-specific variables, but now includes a relationship-specific variable. This variable indicates which types of accounts the firm holds with the bank (deposits and loans, alternatively loans only). The thought behind adding variables stepwise, is to measure whether the new variables are significant, and how they affect the existing variables.

Table 9: Partial likelihood estimates of proportional hazard model (Cox)

	Cox Model			
	1	**2**	**3**	**4**
Ln Income	1.009	0.996	0.993	0.988
	(0.007)	(0.007)	(0.008)	(0.007)
Age At Start	0.998*	0.999	1.000	0.998**
	(0.001)	(0.001)	(0.001)	(0.001)
Profitability	0.855***	0.869***	0.866***	0.861***
	(0.031)	(0.032)	(0.031)	(0.032)
Leverage	0.963***	0.956***	0.954***	0.954***
	(0.011)	(0.011)	(0.011)	(0.012)
Multiple Relationships	1.556***	1.334***	1.254***	1.208***
	(0.054)	(0.068)	(0.064)	(0.062)
Tangibility		0.674***	0.630***	0.541***
		(0.029)	(0.027)	(0.024)
Creditor Concentration		0.489***	0.549***	0.663**
		(0.095)	(0.109)	(0.129)
Nationality			1.962***	1.888***
			(0.075)	(0.072)
Savings Bank			0.394***	0.419***
			(0.017)	(0.018)
Other Bank			1.154***	0.585***
			(0.032)	(0.019)
Deposit & Loan				0.337***
				(0.011)
N	**7,892**	**7,696**	**7,696**	**7,696**

*Significant at 10%, **significant at 5%, ***significant at 1%

The estimates in this table are based on Maximum likelihood estimation of the proportional hazard model using the Cox (1972) partial likelihood function. The hazard ratios measure the partial impact of each variable on the likelihood that a relationship terminates, conditional on duration. Ln Income is the natural logarithm of yearly total income. Age At Start is the years between a firm's founding date and the beginning of the relationship (1998). Profitability is the ratio of operating income to the year-end book value of assets. Leverage is the book value of debt divided by year-end value of total assets. Multiple relations equal one when a firm maintains more than one bank relationship and zero otherwise. Tangibility is fixed assets divided by the year-end total book value of assets. Creditor Concentration is measured by the HHI index and is calculated for each firm by summing the square of loan shares. Nationality equals one when the bank is Norwegian owned and zero otherwise. Finally Deposit & Loan equal one when a firm holds both a deposit and a loan account in the bank and zero if the firm has only a lending relationship with the bank. Hazard ratios are listed in the first row of each cell, with standard errors reported below in parentheses. N is the number of sample observations.

Model 1: Base model

Our base model gives us some evidence to back up the assumption that the selected variables contribute to influencing the duration of a relationship. *Ln Income*, the proxy for size, is insignificant and does not help in explaining the relationship duration. Although *Ln Income* is insignificant, both *Profitability* and *Leverage* are significant with a hazard ratio below one. This is interesting since the low hazard of *Profitability* implies that highly

profitable firms end bank relationships later. The same is the case for the estimate of *Leverage* – the higher the leverage, the longer the relationship. Both these results contradict the studied literature; the former is inconsistent with Ongena and Smith's (1998b) argument that profitable firms are less dependent on bank financing and therefore less susceptible to holdup costs, while the latter is opposed to the findings of Farinha and Santos (2000) that the firms with the greatest need of bank financing are the ones ending the relationship earliest. By relaxing the assumption about significance (and using a 10% level), we can say that duration of a relationship increases with the age of the firm. This is consistent with the hypothesis that firm age increases the likelihood of maintaining a relationship. However, this does not hold for all four of our models. As we shall see, *Age At Start* is insignificant for *Models 2* and *3*. *Multiple Relationships* indisputably affect the length of a relationship; firms with more than one bank relationship maintain shorter relationships. This is consistent with the results of (among others) Farinha and Santos (2000) and confirms our previously stated hypothesis that multiple relationships decrease the length of a relationship. Multiple bank relationships will decrease the value of private information to any one bank, and reduce the ability for any one bank to lock in a customer.

When performing the regression, 7,892 observations are retained. The lost observations are mainly due to missing values in the Dun & Bradstreet database.

Model 2: Base model and extra firm-specifics

In our next model, we keep the same explanatory variables as in the base model, and include other firm-specific variables which we believe may be of relevance for explaining relationship duration. These are *Tangibility* (or the liquidity of the firm) and *Creditor Concentration*. The benefit of having more tangible assets contributes to the likelihood of having a longer bank relationship. This is reasonable, as tangible assets can be seen as firm security or collateral, and is a favorable trait which makes the firm more attractive to the bank, and more likely to achieve lucrative lending conditions.

A firm is more likely to maintain a longer relationship when *Creditor Concentration* is high (i.e. a large Herfindahl-Hirschman index value). High creditor concentration means that a firm tends to concentrate its loans around one source. This is consistent with the findings of

Petersen & Rajan (1994) about firm borrowings; (small) firm borrowings are highly concentrated, even for firms with multiple lending sources.

Another firm-specific characteristic that we included in *Model 2* is *Income Growth*[59]*,* mainly based on the hypothesis that a high growth firm is less dependent on external financing. Hence, the firm may be less dependent on strong bank relationships. In addition Farinha & Santos (2000) find that a high growth firm is expected to borrow a significant amount from new banks in order to promote competition with the existing bank and thereby creating independence of one bank. However, including growth reduces our sample considerably (by about 1500 observations), which may cause problematic results. There is a risk that the variable will become explanatory as a consequence of the smaller sample set, and the problem arises if the smaller sample set is no longer comparable to the original sample. By testing this, we found that the reduced sample (when including income growth) is actually representative for the original *Model 2*.[60] However, because of the large reduction of observations, and to avoid further complications of the model, we decided to subtract growth from our further analysis. Without the growth variable, *Model 2* makes use of a sample selection of 7,686 observations.

Model 3: Extended base model and bank specifics

In Model 3 we still sustain the variables of the base model and the extended firm specifics, and now add bank-specific variables. The bank specifics are nationality of the bank (owner) providing the loan (and deposit), and bank category (i.e. commercial bank, savings bank or other type of bank)[61].

Nationality exhibits whether the bank of the relationship is Norwegian owned or foreign owned with a branch in Norway (i.e. placed under Norwegian tax laws and thus under Norwegian supervision). Our sample has a distinct majority of Norwegian banks (almost 87% of the sample), as shown in the descriptive statistics. Dramatically enough, the estimated hazard ratio shows that the duration of relationships with Norwegian banks seem to be

[59] Calculated by dividing a firm's operating income by the operating income the previous year.

[60] *Income Growth* became significant and likely to shorten the relationship duration. This is consistent with Farinha and Santos' (2000) findings of large borrowings to avoid "lock-in". (See regression result in appendix, Table E.)

[61] Note that since bank category is a *category variable* (and not an indicator variable), it is not directly included in the model. A category variable that takes on *m* distinct values is converted into *m-1* indicators, where one category value is omitted. In our regression results *commercial bank* is omitted.

shorter than those with foreign banks. A potential reason may be the structural changes of the Norwegian bank market mentioned in the descriptive overview of section 3. After 2001 international (primarily Nordic) banks heavily entered the Norwegian market, thereby increasing the competition. In order to gain customers, the foreign banks may initially have offered favorable lending conditions in terms of lower spread, lower monitoring costs, better services provided and so on. This can be seen in relation to the statement in section 2, that banks are willing to make temporary sacrifices to obtain future benefits.

Our data set enables us to make a "quick and dirty" analysis of the assumption that foreign banks offered a lower spread than the Norwegian banks. We know each firm's year-end amount of interest paid, and we know the year-end amount of borrowing. In addition, we have extracted the NIBOR 3-month rate as a reference of the inter-bank rate of interest[62]. One way of calculating the spread for each firm-bank relationship per year will be:

$$spread_t = \frac{interest\ paid_t}{amount\ of\ loan_t} - NIBOR_{3m,t}$$ [63]

However, when generating the spread we get some highly illogical and irrational values (extremely high spread). Even applying 90% Winsorization does not properly adjust the irregularities. One possible explanation for the illogical values is that the data set only contains year-end information. One hypothesis is that firms may refinance or pay off debt at the end of the year; thus the interest rate paid during the year is not corresponding to the actual year-end amount of loan reported (leading to a relatively high numerator in the equation above).

Despite the illogical, inaccurate spread values, we may be able to make some assumptions based on tendencies; the average spread of Norwegian banks in our sample is 7400%, while the corresponding value for international banks is 1600%[64]. This may indicate what we suggested earlier, that international banks offered lower rates than Norwegian banks.

In our sample, there is an evident majority of relationships with commercial and savings banks. Savings banks have longer relationship durations than commercial banks, while the

[62] Norges Bank (Norway's central bank)

[63] A more accurate measure would be to use the average amount of loan over two time periods. However, the average yields similar strange rates and even more missing values.

[64] If the spread calculation were to produce logical values in percentage, this number should generally be very low. For example in September 2005 the market interest rate was 0.39% above the 3 month NIBOR rate.

other banks category implies shorter relationships than both commercial banks and savings banks. The former is consistent with our sample mainly consisting of relatively small, young firms and the assumption that savings banks tend to maintain relationships with mostly small and medium-sized firms. The latter can be seen in relation to our belief that the other types of lenders are firms which generally provide more short-term financing.

As can be seen from Table 9, adding the bank-specific variables has a small effect on the estimates of the other variables. The sample number of observations is not affected by adding the bank specifics and stays the same as for *Model 2*.

Model 4: Extended base model, bank and relationship specifics

The additional variable in *Model 4* is a relationship-specific characteristic, and indicates whether the bank relationship is a deposit and loan relationship, or just a lending relationship.

Firms with both deposits and loans have longer relationships than firms who only maintain loans with their bank. This correlates with the findings of Kim et al. (2004) that well diversified banks (in terms of services provided and client portfolio) are more attractive, and Rajan's (1997) statement that relationships evolve over time and become mutually more beneficial. It is also reasonable, based on the information problems theory, to believe that the costly information asymmetries are reduced when the firm has both deposits and loans in the bank (more information leads to less asymmetry). An important remark when adding the *Deposit & Loan* variable is that the hazard ratio of *other* bank types changes from increasing the hazard to decreasing the hazard of a relationship length. This indicates that it is not so much the type of bank offering the services, but rather the types of services provided which have the greatest impact on the duration of a relationship. A savings bank still implies the longest relationships, but when influenced by the account type, the other bank types also imply longer duration than commercial banks.

Adding the relationship-specific variable affects the significance of the other variables in the model to some extent. The significance of *Creditor Concentration* is reduced to a 5% level, but more interesting is that by adding the variable *Deposit & Loan*, *Age At Start* becomes statistically significant on a 5% significance level. As previously mentioned, this indicates that the duration of a bank relationship increases with the age of the firm. This is consistent with

Petersen and Rajan's (1994) findings that older firms achieve more favorable lending conditions and Ongena and Smith's (1998b) results that duration increases in firm age.

Our tested sample size remains the same as for *Models 2* and *3* with 7,686 observations.

Summary

The estimated hazards are generally stable throughout the testing. The variation of significance for age has only a small effect on the explanation of the model; as the hazard ratio is very close to one, an increase of firm age constitutes very little influence on the overall survival probability. The most influential covariates in explaining the survival probability are bank nationality, accounting type, bank category and tangibility, while the base firm-specific characteristics (in Model 1) only influence to a lesser extent. In addition, the number of observations in our sample is relatively constant in all four models.

Model 4 (including all covariates) suggests that the conditional likelihood of lengthening the duration of a bank relationship increases by a firm being old, highly profitable and highly leveraged, having more tangible assets, holding both deposits and loans, maintaining high creditor concentration and having a relationship with a savings bank. The likelihood of ending a relationship increases when a firm has multiple bank relationships and when the relationship is with a Norwegian (owned) bank. To sum up; high performing, liquid firms with a high debt ratio tend to maintain longer relationships. In addition, holding both deposits and loans with a single foreign bank and high creditor concentration elongate the relationship.

A somewhat remarkable result is that firm size has no significance in explaining the duration of a relationship. This does not correspond with Ongena and Smith's (1998b) results. In addition, both the profitability and leverage estimates project results that oppose those of the majority of the previous literature.

5.2 Results using restricted baseline hazard models

To check the sensitivity of the regression results, we performed regressions with exponential and Weibull distribution specifications of the baseline hazard function presented in Equations (8) and (10) respectively. In the methodology section we explained that with the

Cox model (1972) we do not need to make assumptions about the function of the baseline hazard. According to Kleinbaum and Klein (2005), the Cox model is "robust" in the sense that the hazard ratios and coefficients from the Cox model can closely approximate the results of a parametric model e.g. the Weibull model. Hence the Cox model results are comparable to the results of the Weibull model. By introducing the parametric models, we specify the functional form of the baseline hazard, meaning that we assume that the distribution of the time to relationship termination has a certain form. As we shall see in this section, the estimates on the explanatory variables from the parametric models are very close to the ones reported in the Cox regression results (Tables F and G provide more detailed regression results for the exponential and Weibull model respectively). For each specification, we only report the results for the full model selected (*Model 4*). We maintain the same number of observations for each specification as in the Cox regressions. The table below provides the regression results.

Table 10: Parametric estimation of proportional hazard model (Exponential and Weibull)

		Model		
		Cox (4a)	Exponential (4b)	Weibull (4c)
$\hat{p}$		-	1	1.385+++
		-	-	(0.014)
Intercept		-	0.484***	0.380***
		-	(0.101)	(0.081)
$\hat{\beta}$	**Ln Income**	0.988	0.989	0.982**
		(0.007)	(0.008)	(0.007)
	Age At Start	0.998**	0.998*	0.997**
		(0.001)	(0.001)	(0.001)
	Profitability	0.861***	0.851***	0.826***
		(0.032)	(0.031)	(0.031)
	Leverage	0.954***	0.955***	0.944***
		(0.012)	(0.012)	(0.012)
	Multiple Relationships	1.208***	1.223***	1.272***
		(0.062)	(0.062)	(0.065)
	Tangibility	0.541***	0.542***	0.449***
		(0.024)	(0.023)	(0.020)
	Creditor Concentration	0.663**	0.703*	0.565***
		(0.129)	(0.136)	(0.112)
	Nationality	1.888***	1.849***	2.073***
		(0.072)	(0.070)	(0.079)
	Savings Bank	0.419***	0.428***	0.367***
		(0.018)	(0.018)	(0.016)
	Other Bank	0.585***	0.616***	0.545***
		(0.019)	(0.020)	(0.018)
	Deposit & Loan	0.337***	0.347***	0.273***
		(0.011)	(0.011)	(0.009)
Median duration			5.0	5.3
			(5.7)	(5.0)
N		**7,696**	**7,696**	**7,696**

*Significant at 10%, **significant at 5%, ***significant at 1%
+++, ++, + p=1 can be rejected at 1%, 5% and 10% respectively

In this table we compare the Cox model with the parametric models (exponential and Weibull). The estimates in model 4a are based on maximum likelihood estimation of the proportional hazard model using the Cox (1972) partial likelihood function. The estimates in models 4b and 4c are respectively based on maximum likelihood estimation of the proportional hazard model using the exponential- and Weibull distribution as baseline hazard rates. The parameter, p, measures the degree of duration dependence. The exponential model assumes p =1. The hazard ratios measure the partial impact of each variable on the likelihood that a relationship terminates, conditional on duration. Hazard ratios are listed in the first row of each cell, with standard errors reported below in parentheses. N is the number of sample observations.

The magnitude and significance of the hazard estimates with exponential and Weibull distribution are analogous to those for the Cox regression estimates, indicating that they are only marginally sensitive to the specification of the baseline hazard function. However, there are some differences in the hazards; the hazard of size (*Ln Income*) becomes significant (on a

5% level) in the Weibull regression. *Age At Start* is statistically significant (at a ten percent level) for both specifications. Firm size and age influencing the relationship duration are consistent with the literature stating that the duration increases in firm size and age. *Creditor Concentration* becomes significant at a ten percent level of the exponential regression and a one percent level of the Weibull regression. Note that an estimated hazard of 2.07 for *Nationality* does not mean that a Norwegian owned bank fails twice as fast as an international owned bank. A hazard ratio above two, means that the relationship is likely to fail faster, but in a very specific sense. In the context of hazard ratio, "fast" means (isolated to bank nationality) that a bank relationship with a Norwegian bank which has not yet failed by a certain time has twice the risk of failing at the next point in time compared to a firm maintaining a relationship with an international bank (Grace et al., 2004).

Recalling that the connection between hazards and coefficients can be written as $ln\, t = -\beta_x x_j$ [65], we can also interpret the explanatory variables in terms of their influence on the expected duration of the relationship[66]. For example, the estimate of the intercept in the Weibull regression implies that the expected length of a bank relationship is 2.6 years, when all the explanatory variables are zero.[67] Similarly, the estimated hazard of multiple-bank relationships implies that multiple-bank relationships are expected to last 2.1 years[68].

Recall from the methodology section that while the exponential model assumes constant duration dependence (the parameter *p=1*), the Weibull model allows us to test for the presence of positive duration dependence mentioned in the previous section. Note that the Weibull model estimates $\hat{p}$ to be significantly greater than one (see Table 11 below), implying that bank relationships exhibit positive duration dependence.[69] The Weibull distribution assumes duration dependence to be monotonic over the survival time. This means that the model does not take into account the fact that the likelihood of a relationship ending may increase over time in the earlier part of the relationship (positive

[65] By taking the logarithm on each side of the equation $h(t|x_j) = h_0(t) * e^{x_j \beta_x}$, where the baseline hazard is heuristically removed.

[66] Because duration is inversely related to the hazard rate, a positive (negative) slope coefficient estimate implies a shorter (longer) duration.

[67] Intercept coefficient = - 0.967 → $t = e^{-(-0.967)} = 2.63$

[68] *h(t) = 1.27* → $\hat{\beta} = 0.24$ → $t = e^{-(-0.967+0.24)} = 2.07$

[69] The formal test is a Wald test for H_0: *ln(p) = 0*, which is the equivalent to testing H_0: *p = 1.* The result for all four models that we can reject that the hazard is constant.

duration dependence), but that the likelihood then decreases later in the relationship (negative duration dependence). By looking at the Kaplan Meier plot (Figure 7) it can be argued that the duration dependence is not monotonic. To take this into account it is possible to introduce a third parametric model, the log-logistic. The log-logistic distribution allows for non-monotonic duration dependence.[70] We estimated the four models using log-logistic distribution (although not reported), without gaining new information regarding the hazards of the explanatory variables. Overall, the positive duration dependence essentially supports the idea that firms will be more likely to leave the bank as the relationship lengthens.

[70] The model assumes a baseline hazard of the form: $h_0(t) = \frac{pt^{p-1}\,exp(\beta 0)}{1+t^p\,exp(\beta 0)}$ for p>0, where *p* is some ancillary shape parameter estimated from the data and the scale parameter is parameterized as $exp(\beta_0)$.

5.3 Robustness tests

We shall now examine the robustness of our results by analyzing previously unexamined reasons for observing bank terminations. For all the results we report only the Weibull specification of the full model selected (*Model 4*), as depicted in Table 11. The number of observations is consistent throughout the testing of the additional variables.

Table 11: Robustness tests - Estimations of the Weibull model

		Model			
		Weibull (4c)	Market Concentration (5)	Primary Bank (6)	Loan Amount (7)
$\hat{p}$		1.385+++	1.386***	1.386***	1.394***
		(0.014)	(0.014)	(0.014)	(0.014)
Intercept		0.380***	0.389***	0.387***	0.550***
		(0.081)	(0.088)	(0.083)	(0.119)
$\hat{\beta}$	Ln Income	0.982**	0.982**	0.982**	0.982**
		(0.007)	(0.007)	(0.007)	(0.007)
	Age At Start	0.997**	0.997**	0.997**	0.998**
		(0.001)	(0.001)	(0.001)	(0.001)
	Profitability	0.826***	0.828***	0.824***	0.848***
		(0.031)	(0.031)	(0.031)	(0.031)
	Leverage	0.944***	0.944***	0.944***	0.948***
		(0.012)	(0.012)	(0.012)	(0.012)
	Multiple Relationships	1.272***	1.268***	1.278***	1.282***
		(0.065)	(0.065)	(0.066)	(0.066)
	Tangibility	0.449***	0.450***	0.443***	0.484***
		(0.020)	(0.020)	(0.020)	(0.022)
	Creditor Concentration	0.565***	0.569***	0.538***	0.507***
		(0.112)	(0.113)	(0.108)	(0.100)
	Nationality	2.073***	2.076***	2.070***	2.063***
		(0.079)	(0.080)	(0.079)	(0.079)
	Savings Bank	0.367***	0.367***	0.366***	0.376***
		(0.016)	(0.015)	(0.015)	(0.016)
	Other Bank	0.545***	0.544***	0.538***	0.582***
		(0.018)	(0.018)	(0.018)	(0.020)
	Deposit & Loan	0.273***	0.272***	0.276***	0.262***
		(0.009)	(0.009)	(0.010)	(0.009)
	Added control variable	-	0.746	1.057*	0.979***
		-	(0.605)	(0.034)	(0.002)
Median duration		4.3	5.3	5.3	5.5
		(5.0)	(4.9)	(5.3)	(6.86)
N		7,696	7,686	7,696	7,696

*Significant at 10%, **significant at 5%, ***significant at 1%

+++, ++, + p=1 can be rejected at 1%, 5% and 10% respectively

The estimates in this table are based on maximum likelihood estimation of the proportional hazard model using the Weibull distribution as baseline hazard rates. The parameter, p, measures the degree of duration dependence. Market Concentration is measured by the Herfindahl-Hirschman index (HHI) and estimates the bank market centralization of the region in which the bank is located. This HHI is defined as the sum of squared bank share lending for each region per year. Primary bank equals 1 if the bank in the relationship is lending more than 50% of the total loan taken up by the firm, and zero otherwise. Loan Amount is the amount of loan taken up in the bank (in NOK). Hazard ratios are listed in the first row of each cell, with standard errors reported below in parentheses. N is the number of sample observations.

Market Concentration

As described in section 3, Norway has many strong regional banks throughout the country, which may indicate that regional market concentration is generally low. Petersen and Rajan (1995) find that young firms are more likely to take up bank financing in concentrated credit markets, while older firms borrow less from institutions and are less influenced by the concentration of the market. We add the variable *Market Concentration* to determine if more concentrated bank markets influence the relationship duration where our hypothesis is that higher bank market concentration increases the duration of the relationship, due to few attractive bank alternatives. In such an environment, it is reasonable to argue that it is easier to become "locked in". The HHI bank concentration is defined as the sum of squared bank lending share for each region per year. As shown in Table 11 (appearing in the column labeled "added control variable") the estimator of the additional variable is statistically insignificant, indicating that the bank relationship length is independent of how concentrated the bank market is.[71]

Primary bank

Bharath et al. (2009) study the difference in cost between a lead lender and syndicated loan members. They refer to the "Syndicate Moral Hazard" which occurs as information asymmetries arise among lenders and due to the lead primary bank's incentive to avoid monitoring costs. Bharath et al. (2009) examine the impact of relationships in lowering such information asymmetries between multiple lenders and find that repeated borrowing from the same lender is associated with an almost 10 basis points lower spread on average, lower collateral requirements, shorter maturity of the loan and better access to credit. Hoshi et al. (1991) find that firms with close ties to banks are less likely to suffer from financial distress and conclude that the primary bank relationship plays a role in mitigating the asymmetric information problem. 67% of our total sample consist of relationships with a primary bank. To test the hypothesis that primary banks mitigate the information asymmetry resulting in a stronger bank relationship, we add the variable *Primary Bank* which proxies as the biggest lender (by amount of the borrowings) where the dummy equals one if the amount of the loan is more than half of the total loans taken up by a firm, and zero otherwise. The dummy

[71] Note that the highest bank market concentration can be found in "Nord-Norge" (HHI ratio = 0.13), and the lowest concentration is observed in "Østviken" (HHI =0.09). This shows that the concentration is in general low in Norway during the sample period. A full overview can be found in Table H in the appendix.

becomes significant (on a 10% level) in the regression (*Model 6*), and the significance and magnitude of the original model estimates remain robust to the addition of *Primary Bank*. However, the additional dummy indicates that maintaining a relationship with the primary bank shortens the relationship duration which is contradictive to the findings of Hoshi et al. (1991). In other words, for the firms in our data set, there is little reason to believe that a relationship with a primary bank necessarily means achieving a lower spread, lower collateral and generally better conditions on loans. As Table 11 shows, a relationship with a primary bank implies a 5.7% higher hazard ratio than a non-primary relationship. This may indicate that the primary bank advantage first becomes beneficial for the firm when the relationship matures.

Loan Amount

One of the unique features of our data set is that it contains the amount of borrowings for each firm[72]. It is common to argue that the firms most dependent on external bank funding gain most from strong bank relationships (e.g. small, young firms). Based on this assumption we argue that the size of loans taken up in a bank is also an indication of how dependent the firm is on the relationship. Our hypothesis is that large loans have a positive impact on the duration of bank relationships. The average amount of loan for a firm in our sample is 2.1 million NOK. The variable Loan Amount for each relationship is the quantity of borrowing from each bank.[73] The estimator for *Loan Amount* (shown in *Model 7* in Table 11 above) becomes significant and the hazard ratio indicates that it may contribute to extending the relationship duration, although only to a small extent as the ratio is very close to one. The model remains robust to the additional variable as the magnitude of the other hazard ratios remains more or less unaffected.

[72] Recall from the data section that this is year-end values from the balance sheet.
[73] The variable is scaled by using the natural logarithm of lending.

6. Discussion and conclusion

By utilizing a unique data set of bank relationships from 1998 to 2008, we analyze the determinants that influence whether the relationships are likely to be long-lasting. Our empirical results show that a bank relationship started in 1998 has a median duration of four years, and only 16% of the relationships are ongoing by the end of the sample period. Testing for various firm-, bank- and relationship specifics under the assumption of the Weibull distribution, we document the presence of positive duration dependence in bank relationships. In other words, the distribution has an upward-sloping hazard function and firms become more likely to end a bank relationship as the relationship matures. This suggests that the value of relationships declines over time, that switching costs are not prohibitively large, and that firms try to avoid holdup costs.

We build a model consisting of ten different determinants by first employing the semiparametric Cox partial likelihood model, then expand (and conclude) the analysis by testing the model using the parametric exponential– and Weibull model. The final Weibull model is robust as the estimates are generally stable and only marginally sensitive to the other parametric specifications.

The estimated relationship between the duration of a relationship and the explanatory variables presents the following results (sorted by the degree of influence to the survival probability): The likelihood of ending a bank relationship *decreases* (duration of relationship increases) when a firm maintains a relationship with a savings bank and holds both deposits and loans. We also find that the length of the relationship is longer when a firm has a high tangibility ratio, has high creditor concentration and is highly profitable. To a lesser extent, the leverage ratio, size and age of the firm also decrease the likelihood of early termination. The likelihood of ending a relationship *increases* when the relationship is with a Norwegian (owned) bank and when a firm has multiple bank relationships. The observation that multiple-bank firms are more likely to end a bank relationship is consistent with the information-based arguments that multiple bank relationships will decrease the value of private information to any (one) bank, reducing the ability for any (one) bank to lock in a customer (Ongena and Smith, 1998b). The results are also consistent with the notion that

multiple-bank customers find it less costly to credibly communicate their value to the public, decreasing the value of holding any (one) bank relationship.

The result that savings banks are more likely to maintain longer relationships relative to commercial banks can be seen in the light of the Norwegian bank market consisting of numerous strong regional banks, and that savings banks tend to maintain relationships with the type of firms mainly present in our sample. A more surprising result is that relationships with Norwegian banks tend to have shorter duration than foreign banks. The proxy for bank nationality has a strong influence on the outcome of the model, and when taking the right-censoring issue into account, it is probable that this estimation result could be even stronger. A likely explanation is the structure change of the bank market after 2001, and a vast introduction of mainly Nordic banks on the Norwegian market as a consequence of this change.

We find some remarkable results of the estimated firm characteristics profitability and leverage. Our findings provide empirical evidence which disputes Farinha and Santos' (2000) findings; we document that the firms with the greatest need of bank financing stay in their relationships longest, and are less likely to end a bank relationship than firms where bank financing may be of a secondary need. An argument which supports our findings is that a high-debt financed firm within the bank-dominated Norwegian economy should find it more costly to leave a particular bank relationship if banks maintain monopoly power. In addition, these results may be seen in relation to the characteristics of our sample, as the selected sample contains relatively young, small, leveraged firms – characteristics identifying new businesses. A puzzling result is that the estimates of our base firm-specific characteristics exert such little influence on the length of a relationship. Especially firm size has a much weaker effect than expected. A possible reason is the previously mentioned characteristic sample firm; an average firm is relatively small.

Drawing too many strong conclusions from our study would be unwarranted, since our data set suffers from some shortcomings. We are unable to observe the reason for the termination of a firm-bank relationship. In addition, as we do not have any information about the bank relationships continuing after 2008, the full effect of explanatory variables on rigth-censored cases lacks a thorough analysis.

A possible avenue for future research is to investigate the concentration of rigth-censored observations around certain characteristics, and the actual effect these have on the duration of a relationship. Another unexamined issue is how mergers and acquisitions influence the data set in terms of merged/acquired banks exiting the sample and thus terminating the firm-bank relationship.

7. References

Books, Journals and lecture notes:

Angelini, P., R. Di Salvo and G. Ferri, 1998. *Availability and cost of credit for small businesses: customer relationships and credit cooperatives*. Journal of Banking & Finance 22, p. 925-954.

Berger, A. and G.F. Udell, 1994. *Lines of credit and relationship lending in small firm finance*. Working paper Wharton Financial Institutions Center, p. 94-11.

Berglöf, E. and H. Sjögren, 1995. *Combining Arm's-Length and Control Oriented Finance – Evidence from Main Bank Relationships in Sweden.* Mimeo Stockholm University.

Bhagat, S., N. Moyen and I. Suh, 2005. *Investment and internal funds of distressed firms*. Journal of Corporate Finance 11, p.449-472.

Bharath, S.T., S. Dahiya, A. Saunders and A. Srinivasan, 2009. *Lending relationships and loan contract terms*. Forthcoming publication in Review of Financial Studies.

Bhattacharya, S and A.V. Thakor,1993. *Contemporary banking theory*, J. Finance. Intermed. 3, p. 2-50.

Blackwell, D.W. and D.B. Winters, 1997. *Banking Relationships and the Effect of Monitoring on Loan Pricing.* Journal of Financial Research 20, p. 275-289.

Bris, A., Welch, I., 2005. *The Optimal Concentration of Creditors*. Journal of Finance 60, p. 2193-2212.

Cesarini, F., 1994. *The Relationship between Banks and Firms in Italy: a Banker's View.* Review of Economic Conditions in Italy, p. 29-50.

Cleves, M.A., W.W. Gould and R.G. Gutierrez, 2004. *An Introduction to Survival Analysis Using STATA(Ch.1-13)*. Stata Corporation, Texas.

Cole, R. 1998. *The importance of relationships to the availability of credit*. Journal of Banking & Finance 22, p. 959-977.

D'Auria, C., A. Foglia and P.M. Reedtz, 1999. *Bank interest rates and credit relationships in Italy*. J. Banking Finance 23, p. 1067–1093.

Degryse, H. and P. Van Cayseele, 1998. *Relationship lending within a bank-based system: evidence from European small business data*. Journal of Financial Intermediation 9, p. 90-109.

Degryse, H.A., S. Ongena and G. Tumer-Alkan, 2009. *Lending technology, bank organization and competition.* Journal of Financial Transformation 26, p. 24-30.

Detragiache, E., P. Garella, L. Guiso, 1997. *Multiple versus Single Banking Relationships: Theory and Evidence*. Journal of Finance 55, p. 1133-1161.

Diamond, D., 1984. *Financial intermediation and delegated monitoring*. Review of Economic Studies 51, p. 393-414.

Diamond, D., 1991. *Monitoring and reputation: the choice between bank loans and directly placed debt*. Journal of Political Economy 99, p. 689-721.

Elsas, R. and J.P. Krahnen, 1998. *Is relationship lending special? Evidence from credit-file data in Germany*. Journal of Banking & Finance 22, 1283-1316.

Fama, E., 1985. *What's different about banks?* Journal of Monetary Economics 15, p. 29-39.

Farinha, L.A. and J.A.C. Santos, 2000. *Switching from single to multiple bank lending relationships: determinants and implications.* Unpublished working paper, Bank for International Settlements, Basle.

Fox, J., 2006. *Introduction to Survival Analysis.* Lecture Notes in the course "Sociology" (761), McMaster University, Hamilton, Ontario, Canada [Online]
Available at: http://socserv.mcmaster.ca/jfox/Courses/soc761/survival-analysis.pdf
[accessed on 28 September 2009].

Garson, G.D., 2009. *Kaplan-Meier Survival Analysis*. Lecture notes in the course "Quantitative Research in Public Administration" (PA 765 – 766), North Carolina State University (NCSU), North Carolina, US [Online]. Available at:
http://faculty.chass.ncsu.edu/garson/PA765/kaplanmeier.htm [assessed 12 October 2009].

Grace, M., J.E. Reid, M. Samore and S.L. Spruance, 2004. *Hazard ratio in clinical trials*. Antimicrobial Agents and Chemotherapy 48, 2787-2792.

Greenbaum, S.I., G. Kanatas and I. Venezia, 1989. *Equilibrium loan pricing under the bank-client relationship*. Journal of Banking & Finance, p. 221-235.

Greenwood, M.,1926. *The natural duration of cancer.* Reports on Public Health and Medical Subjects 33, p. 1-26.

Guiso, L. and R. Minetti, 2004. *Multiple creditors and information rights: theory and evidence from US Firms,* CEPR Discussion Paper No. 4278.

Harhoff, D. and T. Körting, 1998. *How Many Creditors Does it Take to Tango?* Wissenschaftszentrum Berlin.

Heckman, J.J. and B. Singer, 1984. *Econometric duration analysis. Journal of Econometrics 24*, p. 63-132,p. 187-201.

Horiuchi, T., F. Packer and S. Fukuda, 1988. *What Role Has the 'Main Bank' Played in Japan?* Journal of Japanese and International Economies 2, p. 159-180.

Horiuchi, T., 1993. *An Empirical Overview of the Japanese Main Bank Relationship in Relation to Firm Size.* Rivista Internationale di Scienze Economiche e Commerciale 40, p. 997-1018.

Horiuchi, T., 1994. *The Effect of Firm Status on Banking Relationships and Loan Syndication.* In M. Aoki and H. Patrick, eds., The Japanese Main Bank System, p. 258-294, Oxford University Press.

Hoshi, T., A. Kashyap and D. Scharfstein, 1991. *Corporate structure, liquidity and investment: evidence from Japanese industrial groups*. Quarterly Journal of Economics 106, p. 33-60.

Houston, J. and C. James, 1996. Bank information monopolies and the mix of private and public debt claims. Journal of Finance 51, p. 1863-1889.

Kabelfleisch, J. D. and R. L.Prentice, 1980. *The Statistical Analysis of Failure Time.* Wiley, New York.

Kaplan, E. L, and M.L. Meier,1958. *Nonparametric estimation from incomplete observations*. Journal of the American Statistical Association 53, p. 457 481.

Kennedy, Peter,2003, *A Guide To Econometric,* 5th edition, p. 287- 315. MPG Books, Bodmin, Cornwall, UK.

Kim, M., Kristiansen, E.G., Vale, B., 2004. *Endogenous product differentiation in credit markets: what do borrowers pay for?* Journal of Banking & Finance 29, p. 681-699.

Kim, M., Kliger, D., Vale, B., 2001. *Estimating switching costs: the case of banking*. Journal of Financial Intermediation 12, p. 25-56.

Kleinbaum, D. G. and M. Klein, 2005. *Survival Analysis (ch. 3).* Springer Sc.+Business Media, Inc.

Masonson, A., 1992. *Treasury Management: Trends in Cash Management Services.* Healthcare Financial Management 46, p. 70-76.

Mjøs, A. and K. Øksnes, 2009. *Dokumentasjon og kvalitetssikring av SNFs og NHHs database med regnskaps- og foretaksinformasjon for norske selskaper*. Working Paper. SNF, Bergen.

Ongena, S. and D.C. Smith, 1998a. *Bank relationships: a review. Performance of Financial Institutions*, p. 221-258.

Ongena, S. and D. C. Smith, 1998b. *The duration of bank relationships*. Journal of Financial Economics 61, p. 449-475.

Ongena, S., G. Tümer-Alkan and N. von Westernhagen, 2007. *Creditor concentration: an empirical investigation*. Discussion Paper Deutsche Bundesbank, 15/2007.

Pagano, M., F. Panetta and L. Zingales, 1998. *Why do companies go public? An empirical analysis.* Journal of Finance 53, p.27-64.

Petersen, M.A. and R.G. Rajan, 1994. *The benefits of lending relationships: evidence from small business data*. Journal of Finance 49, p. 3-37.

Petersen, M.A. and R.G. Rajan, 1995. *The effect of credit market competition on lending relationships*. Quarterly Journal of Economics 110, p. 406-443.

Rajan, R.G., 1992. *Insiders and outsiders: the choice between informed and arm's-length debt*. Journal of Finance 47, p. 1367-1400.

Rajan, R.G., 1997. *Is there a future in banking? Towards a new theory of the commercial bank*. Journal of Applied Corporate Finance.

Rodríguez, G., 2009. *Approaches to Survival Modeling.* Lecture notes in the course Generalized Linear Models (WWS509), Princeton University, New Jersey, US [Online] Available at: http://data.princeton.edu/wws509/notes/c7s3.html [accessed 13 October 2009].

Rossignoli, B. and G. Chesini, 1995. *Multi-Banking and Customer Relationships in the Italian Banking System.* Research Papers in Banking and Finance 17. Institute of European Finance.

Sharpe, S.A., 1990. *Asymmetric information, bank lending and implicit contracts: a stylized model of customer relationships.* Journal of Finance 45, p. 1069-1087.

Sjögren, H., 1994. *Long-Term Financial Contracts in the Bank-Orientated Financial System.* Scandinavian Journal of Management 10, p.315-330.

von Thadden, E-L, 1995. *Long-term contracts, short-term investment and monitoring*. Review of Economic Studies 62, p. 557-575.

Zineldin, M., 1995. *Bank-Company Interactions and Relationships: Some Empirical Evidence.* International Journal of Bank Marketing 13, p.30-40.

Webpages:

Brønnøysund Register Centre (*Brønnøysundregistrene) Organisation Portal [online] Available at:* http://www.brreg.no/ *[accessed October/November, 2009]*

DnBNor, Mergers and acquisitions. [online] Available at: https://www.dnbnor.com/about_the_group/mergers_and_acquisitions/deskmergers_and_acquisitions.html [accessed November 11, 2009]

Norges Bank, *NIBOR*. [online] Available at: http://www.norges-bank.no/templates/article__554486.aspx [accessed November 10, 2009]

Statistics Norway (Statistisk Sentralbyrå, SSB), 2008. *Det svinger i norsk økonomi* [online] Available at: http://www.ssb.no/ssp/utg/200805/13/ [accessed October 20, 2009]

Statistics Norway (Statistisk Sentralbyrå, SSB), 2008. *Bedrifter etter fylke* [online] Available at: http://www.ssb.no/ssp/utg/200805/13/ [accessed October 20, 2009]

Statistics Norway (Statistisk Sentralbyrå, SSB), 2008. *Statistical Yearbook of Norway 2008* [online] Available at: http://www.ssb.no/english/yearbook/ [accessed October 19, 2009]

The Norwegian Financial Services Association, 2008 (Finansnæringens Hovedorganisasjon, FNH) *Markedsandeler – utlån* [online] Available at: http://www.google.no/url?sa=t&source=web&ct=res&cd=1&ved=0CAcQFjAA&url=http%3A%2F%2Fwww.fnh.no%2FPageFiles%2F1533%2FMarkedsandeler%2520-%2520utl%25C3%25A5n%2520til%2520person-%2520og%2520bedriftsmarkedet.xls&ei=0sUDS8iJCMjI-QaJuN2tCA&usg=AFQjCNH_i7j_cXNQfW3rmsMk18P4BZVWzA&sig2=LE-fagWJcrfoP3XzI_XGiQ [accessed October 22, 2009]

The Financial Supervisory Authority of Norway (Kredittilsynet), 2005. *Er det fortsatt plass til et mangfold av banker?* [online] Available at: www.sparebankforeningen.no/asset/139/1/139_1.ppt [accessed October 22, 2009]

The Norwegian Savings Banks Association (Sparebankforeningen), 2001. *Nå kommer de utenlandske bankene* [online] Available at: http://www.sparebankforeningen.no/id/1630.0 [accessed October 20, 2009]

The Norwegian Savings Banks Association (Sparebankforeningen), 2004. *Små banker best for smb-bedriftene* [online] Available at: http://www.grunnfondsbevis.no/id/4147.0 [accessed October 20, 2009]

8. Appendices

List of Figures in the Appendix

Figure A Year-by-year distribution of relationships, by type of bank 64
Figure B Year-by-year distribution of relationships, by type of accounts 65
Figure C Year-by-year distribution of relationships, by bank owner nationality 66
Figure D Concentration of relationships, by firm location 68
Figure E Concentration of relationships, by bank location 1998 68
Figure F Concentration of relationships, by industry 69
Figure G Mergers and Acquisitions related to the DnB Nor group 70
Figure H Profitability 70
Figure I Before winsorizing After winsorizing Leverage 71
Figure J Before winsorizing After winsorizing Age At Start 71
Figure K Multiple- vs. Single Bank Relationships 73
Figure L Bank Category 73
Figure M Account type 73
Figure N Bank Ownership 73

List of Tables in the Appendix

Table A Cumulative distribution of firms ending bank relationships 63
Table B The distribution of the firms by the number of bank relationships 63
Table C The concentration of relationships, by firm age 67
Table D Survivor function (Kaplan and Meier estimator used) 72
Table E Cox model including firm's growth 74
Table F Parametric estimation of proportional hazard model (Exponential) 75
Table G Parametric estimation of proportional hazard model (Weibull) 76
Table H Bank Market Concentration, by region 77

8.1 Data overview

Table A
Cumulative distribution of firms ending bank relationships

Year	Firms ending bank relationship, cumulative	% marginal
1998	1,781	18.8 %
1999	3,499	36.9 %
2000	4,527	47.8 %
2001	6,082	64.2 %
2002	6,593	69.6 %
2003	7,083	74.7 %
2004	7,375	77.8 %
2005	7,560	79.8 %
2006	7,816	82.5 %
2007	7,992	84.3 %

This table lists cumulatively by year, the total number and percentage of unique firms that are ending one or more bank relationships. A bank relationship is defined as a relationship between a firm and a bank. In all of the relationships the bank is providing loan(s) to the firm. We define a firm as ending a relationship when the relationship is no longer reported to the tax authorities. All numbers are obtained from the Norwegian Ministry of Finance.

Table B
B1: The distribution of the firms by the number of bank relationships maintained in 1998

Number of bank relationships	Number of firms	%, marginal	%, cumulative
1	8,471	89.39%	89.39%
2	879	9.28%	98.67%
3	93	0.98%	99.65%
4	29	0.31%	99.96%
5	1	0.01%	99.97%
6	2	0.02%	99.99%
9	1	0.01%	100.00%

The three first columns of this table list the distribution of the sample firms by the number of relationships in 1998, when all the relationships in the sample started. The last column provides the cumulative distribution. The number of firms in the sample in 1998 is 9,476. All numbers are obtained from the Norwegian Ministry of Finance.

B2: The distribution of the firms by the number of bank relationships maintained in 2008

Number of bank relationships	Number of firms	%, marginal	%, cumulative
1	1,573	98.31%	98.31%
2	25	1.56%	99.88%
3	2	0.13%	100.00%

The first three columns of this table list the distribution of the sample firms by the number of relationships in 2008, by the end of the sample period. The last column provides the cumulative distribution. The number of firms in the sample in 2008 is 1600. All numbers are obtained from the Norwegian Ministry of Finance.

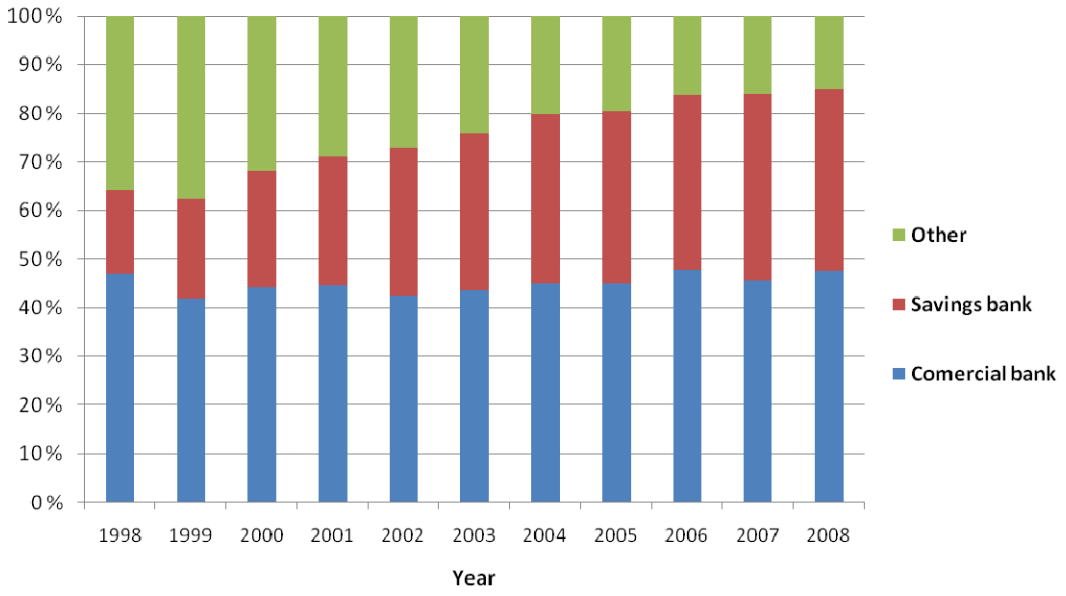


Figure A
Year-by-year distribution of relationships, by type of bank

This figure shows the development of the distribution of relationships by the type of bank from 1998 to 2008. The bank type is defined as by the Norwegian Ministry of Finance, where other is dominated by state/municipal banks, insurance firms and credit firms.

The distribution of relationships, by type of bank in 1998

Bank Category	Number of bank relationships	%, marginal	%, cumulative
Commercial bank	4,750	46.89 %	46.89 %
Savings bank	1,765	17.42 %	64.31 %
Other	3,615	35.69 %	100.00 %

The bank type is defined by the Norwegian Ministry of Finance

The distribution of relationships, by type of bank in 2008

Bank Category	Number of bank relationships	%, marginal	%, cumulative
Commercial bank	766	47.49 %	47.49 %
Savings bank	605	37.51 %	85.00 %
Other	242	15.00 %	100.00 %

The bank type is defined by the Norwegian Ministry of Finance

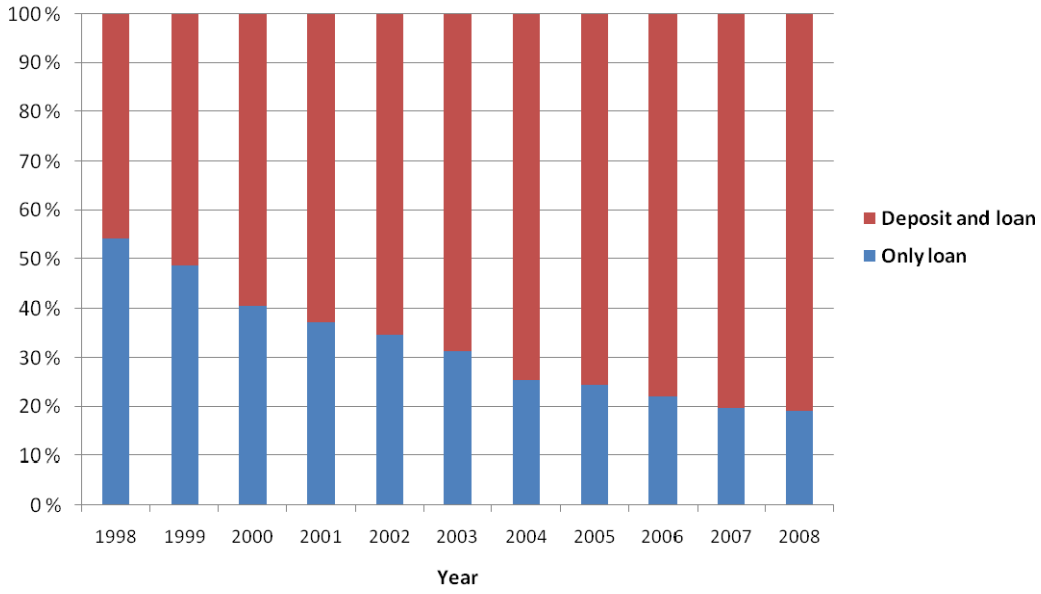


Figure B
Year-by-year distribution of relationships, by type of accounts

This figure shows the development of the distribution of relationships by the type of account the firm had in the bank from 1998 to 2008. In the relationships categorized as "only loan", the firm does not have any deposits throughout the whole sample period

The distribution of relationships, by type of accounts in 1998

Relationship type	Number of bank relationships	%, marginal	%, cumulative
Only loan	5497	54.26%	54.26%
Deposit and loan	4633	45.74%	100.00%

The distribution of relationships, by type of accounts in 2008

Relationship type	Number of bank relationships	%, marginal	%, cumulative
Only loan	310	19.22%	19.22%
Deposit and loan	1303	80.78%	100.00%

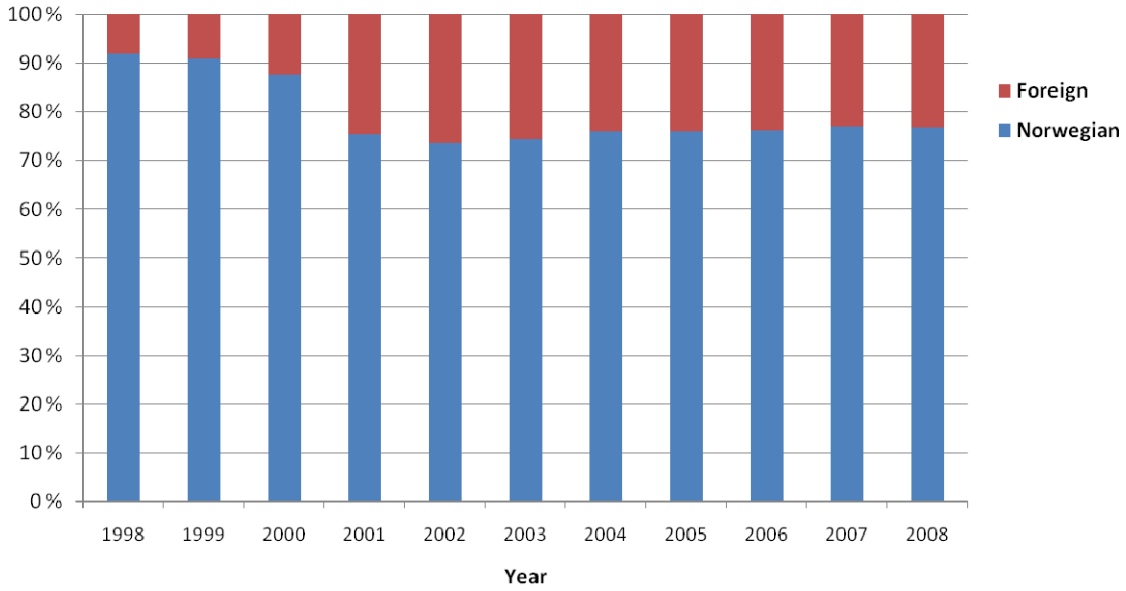


Figure C
Year-by-year distribution of relationships, by bank owner nationality

This figure shows the development of the distribution of relationships by bank owner nationality from 1998 to 2008. We have assumed that the nationality of a bank owner is the same throughout the period.

The distribution of relationships, by bank owner nationality in 1998

Nationality bank owner	Number of bank relationships	%, marginal	%, cumulative
Norwegian	9,310	91.91%	91.91%
Foreign	820	8.09%	100.00%

We have assumed that the nationality of a bank owner is the same throughout the period

The distribution of relationships, by bank owner nationality in 2008

Nationality bank owner	Number of bank relationships	%, marginal	%, cumulative
Norwegian	1,240	76.88%	76.88%
Foreign	373	23.12%	100.00%

We have assumed that the nationality of a bank owner is the same throughout the period

Table C
C1: The concentration of relationships, by firm age in 1998

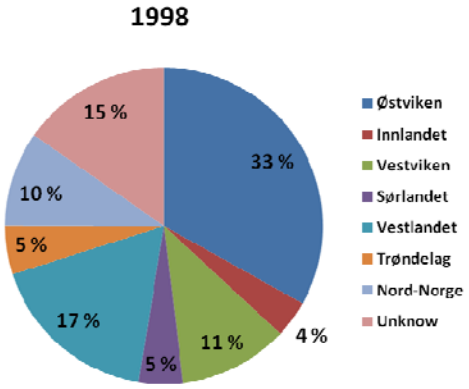
Age	Number of firms	%, marginal	%, cumulative
0 to 5	3,922	41.39%	41.39%
6 to 10	2,304	24.31%	65.70%
11 to 20	1,508	15.91%	81.62%
21 to 40	621	6.55%	88.17%
41 to 70	175	1.85%	90.02%
70+	83	0.88%	90.89%
Unknown	863	9.11%	100.00%

This table shows the distribution of relationships by firm age in 1998. The age in 1998 is defined to be the number of years from the firm was founded (according to the registration in the Brønnøysund Register) till 1998. All numbers are obtained from the Dun & Bradstreet database.

C2: The concentration of relationships, by firm age in 2008

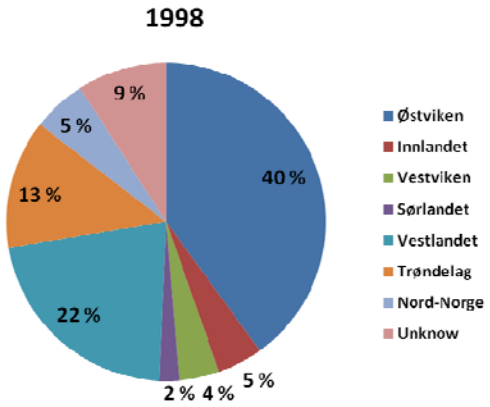
Age	Number of firms	%, marginal	%, cumulative
10 to 15	608	38.00%	38.00%
16 to 20	417	26.06%	64.06%
21 to 30	261	16.31%	80.38%
31 to 50	117	7.31%	87.69%
51 to 80	39	2.44%	90.13%
80+	13	0.81%	90.94%
Unknown	145	9.06%	100.00%

This table shows the distribution of relationships by firm age in 2008. Note that all age intervals are shifted 10 years in this table compared to the one for 1998, due to the fact that all firms related to the relationships that are still ongoing in 2008, naturally are 10 years older. The distribution in the corresponding rows of the two tables, C1 and C2, are comparable.



The info is obtained from the Dun & Bradstreet database.

Figure D
Concentration of relationships, by firm location



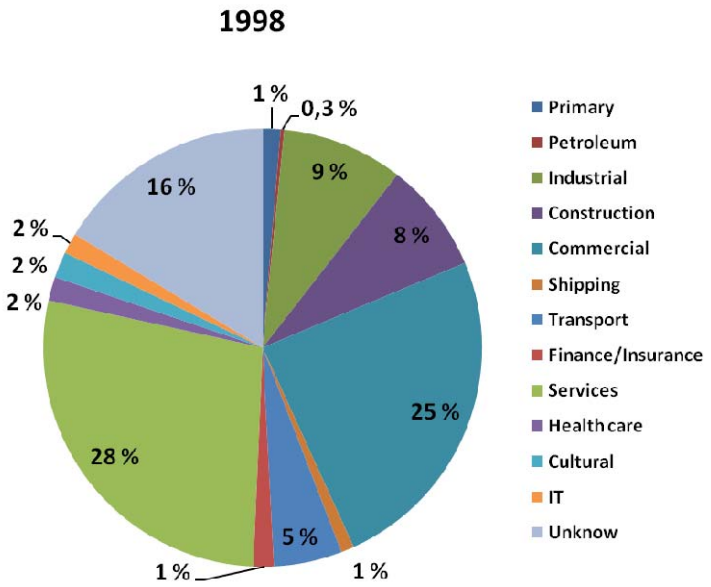
The info is obtained from the Dun & Bradstreet database.

Figure E
Concentration of relationships, by bank location 1998

Provinces related to the seven parts of Norway

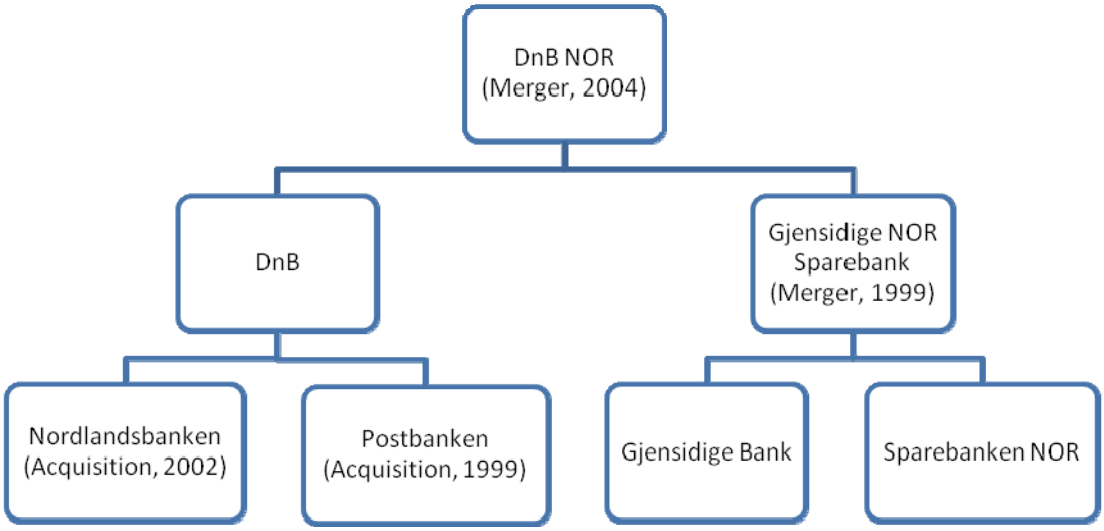
Part of the country	Provinces
Østviken	Østfold, Oslo and Akershus
Innlandet	Hedmark and Oppland
Vestviken	Buskerud, Vestfold and Telemark
Sørlandet	Aust-Agder and Vest-Agder
Vestlandet	Rogaland, Hordaland, Sogn og Fjordane and Møre og Romsdal
Trøndelag	Sør-Trøndelag and Nord-Trøndelag
Nord-Norge	Nordlans, Troms and Finnmark

Statistics Norway (SBB) divides Norway into 7 geographical parts. This table provides information about which provinces (fylker) that are related to each of the 7 parts. The majority of Norwegian banks only operate in one or two of the total 19 counties. Only the three largest banks are represented across the whole country (Kim et al., 2004).



The info is obtained from the Dun & Bradstreet database.

Figure F
Concentration of relationships, by industry

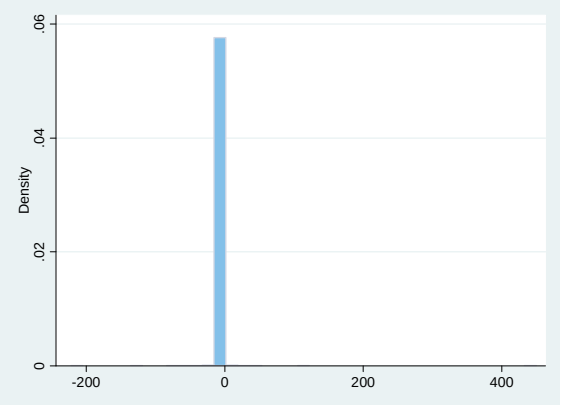


Source: DnB Nor (www.dnbnor.com)

Figure G
Mergers and Acquisitions related to the DnB Nor group

Winsorizing effects

Before winsorizing



After winsorizing

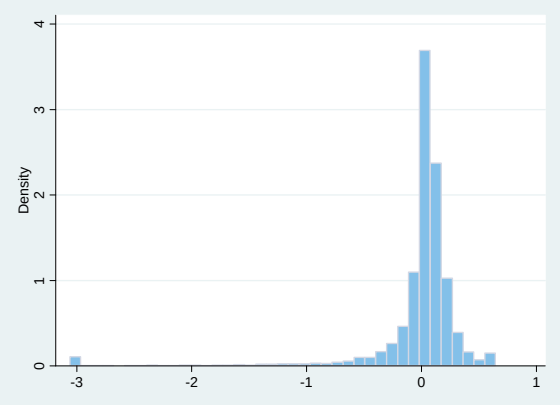


Figure H
Profitability

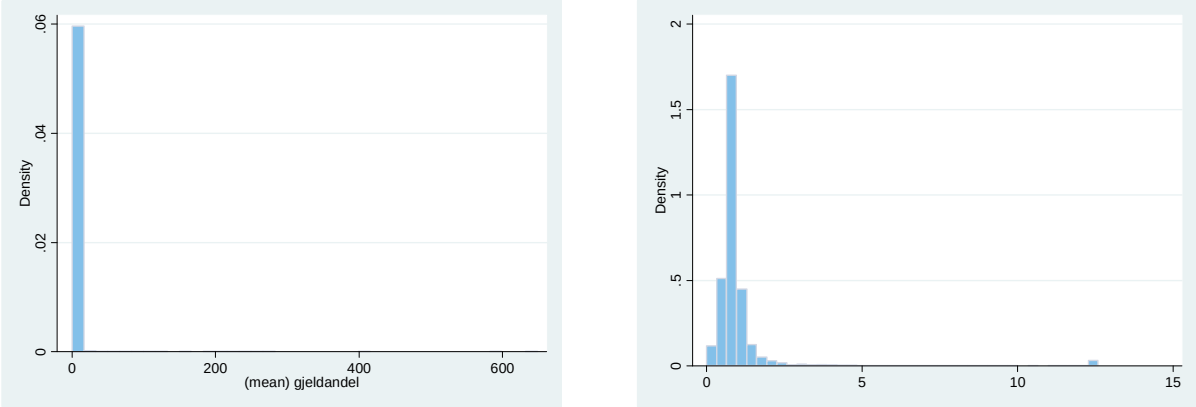


Figure I Before winsorizing After winsorizing
Leverage

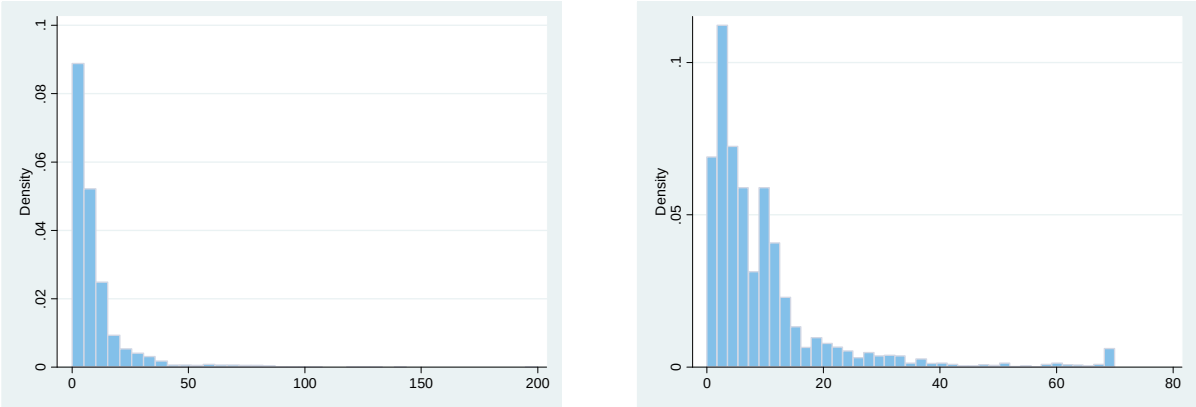


Figure J Before winsorizing After winsorizing
Age At Start

8.2 Econometric definitions

Hazards: The "hazard" is the event of interest occurring. We also refer to it as a failure, and in our study it means that a relationship ends.

Hazard rates: A hazard rate at a given time is the probability of the given event (failure) occurring in that time period, given survival through all prior time intervals.

Hazard ratios: A hazard ratio, also called the hazard function, is the estimate of the ratio of the hazard rate in one group (ex., the treatment group) to the hazard rate in another group (ex. , the placebo group) for a coded covariate (ex., placebo=0, treatment=1). For continues covariates, the hazard ratio is the ratio of the hazard rate given a one unit increase in the covariate to the hazard rate without such an increase.

Source: Cleves, M.A., W.W. Gould and R.G. Gutierrez, 2004. An Introduction to Survival Analysis Using STATA (Ch 2).

8.3 Results

Table D
Survivor function (Kaplan and Meier estimator used)

Year, t	Survivor Function, S(t)	Std. Error	[95% Conf. Int.]	
1998	0.82	0.004	0.81	0.83
1999	0.64	0.005	0.63	0.65
2000	0.53	0.005	0.52	0.54
2001	0.37	0.005	0.36	0.38
2002	0.31	0.005	0.30	0.32
2003	0.26	0.004	0.25	0.27
2004	0.23	0.004	0.22	0.24
2005	0.21	0.004	0.20	0.22
2006	0.18	0.004	0.17	0.19
2007	0.16	0.004	0.15	0.17
2008*	0.16	0.004	0.15	0.17

*No failures observed because of right censoring

This table lists the survivor function, S(t) - the probability of continue the relationship past year t, its annual standard errors and 95% confidence intervals.

Estimated survivor functions by different subgroups

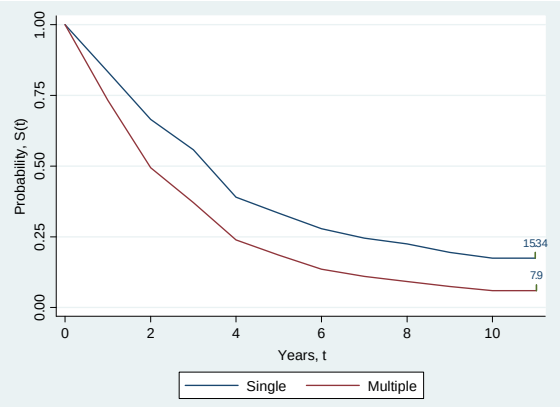


Figure K: Multiple- vs. Single Bank Relationships

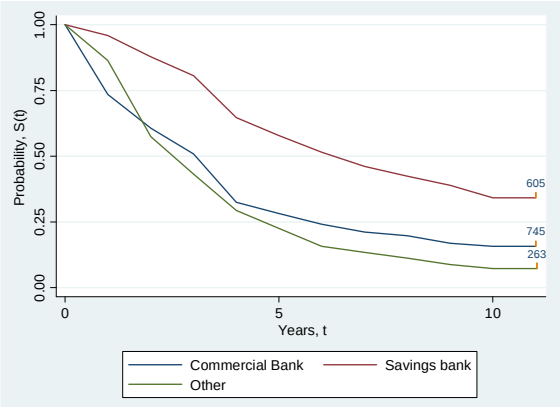


Figure L: Bank Category

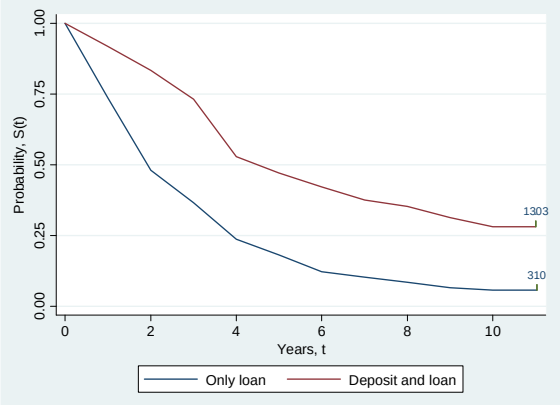


Figure M: Account type

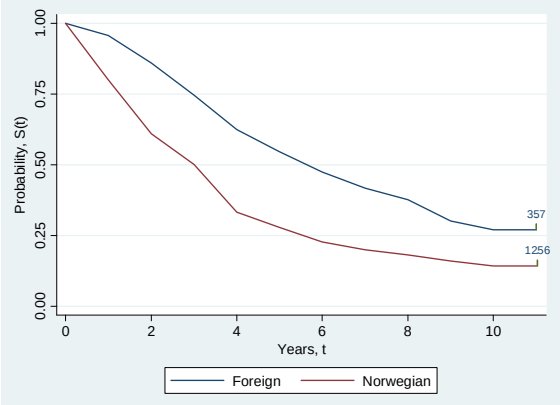


Figure N: Bank Ownership

Number of observations: 10,130. 8,808 (87%) are single bank relationships. 47%, 17% and 35% of the relationships are related to commercial banks, savings banks and other lenders respectively. In 54% of the relationships the firm has only a loan account in the bank, and finally 87% of the relationships are associated with a Norwegian owned bank. The numbers at the end of each curve displays the number of right-censored relationships. For example 1,534 of the censored cases are singe bank relationships while 79 are multiple bank relationships.

Log-rank and Wilcoxon tests

	Log Rank	Wilcoxon
Mult. Vs. Single	237.8***	219.9***
Bank Category	783.1***	803.2***
Account type	1724.8***	1711.1***
Bank Ownership	288.3***	405.89***

The table shows the values of the test estimator (chi2).
* Significant at 10 %, ** significant at 5 %, *** Significant at 1 %

Both the log-rank test and the Wilcoxon test are nonparametric test, allowing us to formally test the equality of survivor functions across two or more groups. They compare the overall survivor functions, under the null hypothesis that the functions of the groups are the same: $H_0=h_1(t)=h_2(t)=\ldots h_r(t)$, where r is the number of groups. The two tests are similar, but the Wilcoxon test places extra weight to earlier failure times (when more observations are at risk).

Table E
Cox model including firm's growth

		Model included Income Growth
$\hat{p}$		1.468+++
		(0.018)
Intercept		0.165***
		(0.042)
$\hat{\beta}$	**Ln Income**	1.010
		(0.009)
	Age At Start	0.998
		(0.001)
	Profitability	0.792***
		(0.034)
	Leverage	0.956***
		(0.013)
	Multiple Relationships	1.468***
		(0.090)
	Tangibility	0.601***
		(0.030)
	Creditor Concentration	0.353***
		(0.086)
	Income Growth	1.021***
		(0.005)
Median duration		6.0
		(1.4)
N		**6,090**

*Significant at 10%, **significant at 5%, ***significant at 1%
+++, ++, + p=1 can be rejected at 1%, 5% and 10% respectively

The estimates in this table are based on maximum likelihood estimation of the proportional hazard model using the Cox (1972) partial likelihood function. The hazard ratios measure the partial impact of each variable on the likelihood a relationship terminates, conditional on duration. In addition to the variable in Model 2, we have included Income Growth, which measures firm's income growth. The variable is defined by dividing a firms (gross) operating income by its operating income the year before. Hazard ratios are listed on the first row in each cell, with standard errors reported below in parentheses. N is the number of sample observations. Note that this variable causes a significant reduction in the sample size.

Table F
Parametric estimation of proportional hazard model (Exponential)

		Model			
		1	**2**	**3**	**4**
$\hat{p}$		1 -	1 -	1 -	1 -
Intercept		0.149*** (0.009)	0.454*** (0.094)	0.248*** (0.053)	0.484*** (0.101)
$\hat{\beta}$	**Ln Income**	1.011 (0.007)	0.996 (0.007)	0.993 (0.008)	0.989 (0.008)
	Age At Start	0.998** (0.001)	0.999 (0.001)	1.000 (0.001)	0.998* (0.001)
	Profitability	0.829*** (0.030)	0.846*** (0.031)	0.854*** (0.031)	0.851*** (0.031)
	Leverage	0.961*** (0.011)	0.953*** (0.011)	0.953*** (0.011)	0.955*** (0.012)
	Multiple Relationships	1.641*** (0.057)	1.388*** (0.071)	1.277*** (0.066)	1.223*** (0.062)
	Tangibility		0.647*** (0.028)	0.609*** (0.026)	0.542*** (0.023)
	Creditor Concentration		0.451*** (0.088)	0.550*** (0.109)	0.703* (0.136)
	Nationality			1.984*** (0.075)	1.849*** (0.070)
	Savings Bank			0.386*** (0.016)	0.428*** (0.018)
	Other Bank			1.194*** (0.033)	0.616*** (0.020)
	Deposit & Loan				0.347*** (0.011)
Median duration		4.3 (0.7)	4.4 (1.6)	4.7 (5.2)	5.0 (5.7)
N		**7,892**	**7,696**	**7,696**	**7,696**

*Significant at 10%, **significant at 5%, ***significant at 1%

The estimates in this table are based on maximum likelihood estimation of the proportional hazard model using the exponential distribution as baseline hazard rates. The parameter, p, measures the degree of duration dependence. The exponential model assumes p =1. Hazard ratios are listed on the first row in each cell, with standard errors reported below in parentheses. N is the number of sample observations.

Table G
Parametric estimation of proportional hazard model (Weibull)

		Model			
		1	2	3	4
$\hat{p}$		1.181+++	1.194+++	1.282+++	1.385+++
		(0.012)	(0.012)	(0.013)	(0.014)
Intercept		0.106***	0.376***	0.171***	0.380***
		(0.007)	(0.078)	(0.037)	(0.081)
$\hat{\beta}$	**Ln Income**	1.012	0.995	0.990	0.982**
		(0.007)	(0.007)	(0.008)	(0.007)
	Age At Start	0.997**	0.999	0.999	0.997**
		(0.001)	(0.001)	(0.001)	(0.001)
	Profitability	0.809***	0.827***	0.829***	0.826***
		(0.030)	(0.030)	(0.030)	(0.031)
	Leverage	0.957***	0.948***	0.945***	0.944***
		(0.011)	(0.011)	(0.011)	(0.012)
	Multiple Relationships	1.740***	1.441***	1.331***	1.272***
		(0.061)	(0.074)	(0.069)	(0.065)
	Tangibility		0.617***	0.551***	0.449***
			(0.027)	(0.024)	(0.020)
	Creditor Concentration		0.390***	0.474***	0.565***
			(0.077)	(0.096)	(0.112)
	Nationality			2.160***	2.073***
				(0.083)	(0.079)
	Savings Bank			0.347***	0.367***
				(0.015)	(0.016)
	Other Bank			1.245***	0.545***
				(0.034)	(0.018)
	Deposit & Loan				0.273***
					(0.009)
Median duration		4.6	4.7	5.0	5.3
		(0.7)	(1.4)	(4.3)	(5.0)
N		**7,892**	**7,696**	**7,696**	**7,696**

*Significant at 10%, **significant at 5%, ***significant at 1%
+++, ++, + p=1 can be rejected at 1%, 5% and 10% respectively

The estimates in this table are based on maximum likelihood estimation of the proportional hazard model using the Weibull distribution as baseline hazard rates. The parameter, p, measures the degree of duration dependence. Hazard ratios are listed on the first row in each cell, with standard errors reported below in parentheses. N is the number of sample observations.

Table H
Bank Market Concentration, by region

Country part	Market Concentration (HHI)
Østviken	0.09
Innlandet	0.09
Vestviken	0.10
Sørlandet	0.11
Vestlandet	0.11
Trøndelag	0.12
Nord-Norge	0.13

The table lists the bank market concentration (HHI) in different regions in Norway. The concentration is measured by the sum of squared bank lending share for each region per year. Note that the table depicts, for each country part, the mean value of the annual HHI ratios from 1998-2008.