Organizational geriatrics: The fittest survive, but they wear out*

Claudio Loderer[†] Klaus Neusser[‡] Urs Waelchli[§]

May 20, 2008

—Preliminary—

Abstract

As they grow old, firms learn to do things better. At the same time, however, they lose their flexibility and their endowments decay and become obsolete. The net effect is unclear. This paper investigates how age affects the chances of survival of firms and their financial performance. The results show that survival is a curvilinear function of age. It goes up at younger ages, and then it declines. Old age eventually leads to organizational death. The results also imply that performance gets worse with age. Better firms survive, but they deteriorate as they become older.

JEL classification: G30, L10

Keywords: firm performance, age, survival, hazard, decay

*We wish to thank Matthias Flückiger for excellent research assistance. Comments from the participants at the XLII EWGFM meeting in Stockholm are gratefully acknowledged. All errors are ours.

[†]Institut für Finanzmanagement, Universität Bern, Engehaldenstrasse 4, CH-3012 Bern, Switzerland. email: claudio.loderer@ifm.unibe.ch

[‡]Department of Economics, University of Bern, Schanzeneckstrasse 1, P.O. Box 8573, CH-3001 Bern, Switzerland. email: klaus.neusser@vwi.unibe.ch

[§]Institut für Finanzmanagement, Universität Bern, Engehaldenstrasse 4, CH-3012 Bern, Switzerland. email: urs.waelchli@ifm.unibe.ch

1 Introduction

Over time, firms learn to be more efficient. They discover ways to standardize, coordinate, and speed up their production processes to reduce costs and improve quality. Much of that activity leads to increased specialization along the lines described by Smith (1776 [1965]) and discussed by David Ricardo (1817). However, as they get older and learn from experience (Arrow (1962)), firms would seem to also encourage structural and process-related rigidities that threaten their chances of survival. The best producers of buggy whips were so focused on their craft that they probably became unable to do anything else. Aging also renders knowledge, abilities, and skills obsolete and induces organizational decay. On balance, it is unclear whether aging helps firms prosper or whether it dooms them. This paper addresses that question.

We investigate two age-related issues, namely survival and performance. The evidence seems to suggest that new firms have a hard time surviving. According to Bartelsman et al. (2005), for example, between 20 and 40 percent of entering firms across ten OECD countries fail within the first two years of life. And only about 40-50 percent of total entering firms survive beyond the seventh year. It would seem that competitive markets for goods and services sweep away inefficient firms. If so, only the fittest survive. The problem is that survival threats do not materialize only in the market for goods and services but also in the market for corporate control. And firms get taken over for various reasons—even when, or perhaps because, they are the best at what they do. If so, surviving organizations might be financially solid but unspectacular.

Various strands of the empirical literature in economics and management are relevant to our research, but few if any empirical papers have looked for answers to the question we are asking. In the management literature, Leonard-Barton (1992) has pointed out that, when firms focus on core capabilities, they bring on core rigidities that make it difficult to adapt to changes in their environment. We argue that these rigidities are compounded by age.

The industrial organization literature has focused on the issues of innovation, entry, exit, and survival. Yet, according to Caves (1998, p. 1958), little attention has been given to organizational geriatrics. One possible reason is that: "The prevalent decline of hazard rates with age suggests that geriatric problems are not serious for firms". Consistent with that assessment, Dunne et al. (1989), study U.S. data from 1972 to 1987 and report a positive relation between firm age and survival. Agarwal and Gort (1996, 2002), however, find traces of a roughly U-shaped relation between age and hazard rates.

We examine how age affects the chances of survival and performance. Our survival analysis extends the work in Agarwal and Gort (1996, 2002) by focusing on the possibly U-shaped relation their data suggest. We also examine the relation between age and performance. The investigation covers all the firms with data on the Compustat and the Compustat Industry Segment database between 1978 and 2004. We define survival as the ability of a given firm to remain an independent organization. Age is the time since listing, as in Clementi (2002).

The paper makes three different contributions to the literature. First, we find that the probability of survival, conditional on a number of variables including size, first increases with age and then declines. Unlike in Agarwal and Gort (1996, 2002), we derive this result regardless of industry and product life cycles. Second, we show that, conditional on age as well as on other variables, better firms are more likely to survive as independent organizations—the fittest survive. Baker and Kennedy (2002)'s results that surviving firms have superior stock price performance are consistent with our finding. Consistent with what the literature on corporate focus seems to suggest, specialization reduces a firm's life expectancy.

Finding that firms which perform better have a better chance to survive, however, doesn't tell us much about the relation between age and performance. It could be that age helps firms get better and that better firms survive. It could also be, however, that age weakens performance, though only the best survive. Daily observations of living organisms support both hypotheses. Young birds, for example, become gradually better at flying—and those that don't probably perish. As age progresses, however, flying skills diminish—and only the better survive. The third contribution this paper makes is to show that getting older is a handicap, even at a young age. It seems that investments and learning cannot overcome the deleterious effects of obsolescence. Older firms perform significantly worse. Better firms survive, but their abilities deteriorate as they get older.

The remainder of the paper is organized as follows. Section 2 presents theoretical considerations to structure the empirical analysis. Section 3 discusses the data of our empirical investigation. Section 4 examines the survival characteristics of our sample firms as a function of various covariates including age. Section 5 inquires into the relation between age and firm performance. Section 6 concludes.

2 Theoretical considerations

Recent papers in the industrial economics literature can help us structure the investigation of the relation between age, survival, and performance. Even though the relevance of age for firm dynamics has attracted comparatively little attention in the literature, several papers have documented or discussed its effects. In particular, it is well-established that there is a positive correlation between survival and age (see, for example, Dunne et al. (1989)). The nature of that observation, and its form, however, are debatable. As pointed out in Thompson (2005), theories that could in principle explain the positive correlation include learning (older firms have more knowledge), financial frictions (Cooley and Quadrini (2001)), and that older firms are active in a larger number of submarkets. Thompson, however, also argues that the observed correlation could simply reflect selection bias. If firms were heterogeneous in their exit hazards then, as a cohort of firms ages, it will become increasingly composed of firms with the lowest propensity to exit, even if the exit hazard does not decline with age for any individual firm.

The evidence in Agarwal and Gort (1996, 2002) suggests that the relation between age and survival might not be monotonous. Hazard rates might pick up with age. They see survival as the trade-off between obsolescence of a firm's original endowments, on the one hand, and net investments and learning-by-doing, on the other. All else being the same, survival depends on three sets of variables: (i) unobserved but age-dependent variables (stock of endowments and learning-by-doing); (ii) observed variables that relate to the attributes of the firm (firm size, as a proxy for incomplete passive learning, and diversification, as reflecting knowledge available for transfer from one market to another); and (iii) observed industry variables (a technology index that affects the rate of obsolescence of the firm's endowments, exit barriers, as measured by factor mobility in the industry, and demand volatility). Learning-by-doing takes two forms. Firms learn about how to do things better (active learning, as in Ericson and Pakes (1995)). Firms, however, also learn about themselves, what tasks they are best suited in carrying out and what markets are the most promising for their abilities (passive learning, as in Jovanovic (1982)). The increase in firm endowments depends on new investment, on the preceding period's stock of endowments, and on accumulated learning. Eventually, the increase in endowments falls below the obsolescence rate. One reason is that the stock of learning increases at a decreasing rate (important lessons are learned first and there is a finite stock of information to be learned about a technology). Another reason is that the adaptability of old endowment diminishes and investment opportunities in new technology shrink as the product market ages.

Our survival analysis takes these relations into account. In the framework of a proportional hazard model (Wooldridge (2002)) we can write:

$$\lambda(t;x) = \kappa(x)\lambda_0(t),$$

where $\kappa(x)$ is a nonnegative function of the vector of explanatory variables x, and $\lambda_0(t)$ is the baseline hazard. $\kappa(x)$ is parameterized as $\kappa(x) = exp(x\beta)$, where β is a parameter vector. According to the preceding arguments, the explanatory variables and the sign of their coefficients are as follows:

Explanatory variables	Expected sign
Age	–
Age^2	+
Initial firm size	– (declining in time)
Current firm size	–
Diversification	–
Hi-tech index	+
Demand volatility	+
Profitability	–
Leverage	+

As mentioned above, age should reduce the exit hazard initially. After a number of years, however, age should bring about obsolescence of capital and skills, and thereby raise the firm's hazard. Hence, the squared term of age should have a positive coefficient. Moreover, firm size, current and past, should increase the firm's endowments and therefore lower the exit hazard. More generous initial endowments should also improve the survival odds of firms, even though their importance should decline over time in an active learning framework (Ericson and Pakes (1995)). By the same argument, capital expenditures should contribute to longer corporate life. The exit hazard increases, however, when firms experience fast obsolescence, such as in high-tech industries. Demand volatility would also seem to threaten survival. Finally, profitability should lower the exit hazard. Since large financial leverage is a threat to survival, we include it among the control variables with a positive sign.

Our analysis differs from that of Agarwal and Gort in that we focus explicitly on the role of age and look for evidence that the exit hazard eventually picks up with age. Moreover, we study the importance of profitability.

The literature has also explored the relation between growth, age, and (especially) size. Evans (1987), for example, finds that firm growth decreases with firm age, and that it decreases at a diminishing rate with firm size. Cooley and Quadrini (2001) develop a model able to predict theses dynamics. The intuition behind their model is that, with financial frictions, firms do not raise all the funds necessary for the marginal product of capital to equal its opportunity cost; consequently, as capital increases over time, its marginal product declines, and so does the firm's rate of growth (Clementi (2002)). Clementi (2002) combines the industrial organization literature on firm dynamics and the corporate finance literature on IPOs. The model he proposes embeds the IPO decision in a dynamic optimization model similar to the one in Cooley and Quadrini (2001). Clementi is able to predict the post-IPO decline in operating return on assets documented in the IPO literature (Jain and Kini (1994) and Mikkelson et al. (1997)). According to the model, firm profitability should be a nonlinear, declining function of age, defined as the number of years after a firm's IPO. We test that prediction in our performance analysis. Profitability is also related to corporate diversification. The literature finds that related diversification improves performance (Villalonga (2004)), whereas unrelated, or conglomerate, diversification impairs it (Lang and Stulz (1994), Berger and Ofek (1995), Servaes (1996), Campa and Kedia (2002)). In our analysis, we therefore control for diversification—unrelated, since the data we use are from COMPUSTAT. In addition, we control for a set of variables similar to that chosen by Campa and Kedia (2002) in their investigation of the diversification discount. We therefore test the following specification of the regression of performance on age and other determining factors:

Explanatory variables	Expected sign
Age	−
Age^2	+
Firm size	−
Diversification	−
CAPEX/Sales	+
Demand volatility	?
Firm size (1 lag)	?
CAPEX/Sales (1 lag)	?

In words, we expect a negative, though nonlinear relation between age and performance, as predicted by Clementi (2002) and Cooley and Quadrini

(2001). Performance should also be negatively related with firm size (Cooley and Quadrini (2001)), and so should conglomerate diversification (as per the literature on the diversification discount). The remaining regression arguments cannot be signed a priori.

3 Data

3.1 Sample description

The sample consists of all firms with data on the COMPUSTAT and the COMPUSTAT Industry Segment database between 1978 and 2004. Following Berger and Ofek (1995), among others, we exclude firm-years with total sales of less than USD 20 million (COMPUSTAT DATA12), firm-years with missing values for total assets (DATA12), and firm-years for which the sum of segment sales deviates from total sales by more than one percent. Unlike other studies, however, we do not exclude firms with business segments in the financial sector (SIC 6000-6999). Our final sample consists of 10,930 firms and 82,845 firm-years, including 1,669 firms (6,644 firm-years) that operate in the financial sector.

Table 1 reports the composition of the sample as well as the number of firms that enter and exit the sample during the 27 years under investigation. We start with 2,285 firms in 1978 and end with 2,923 firms in 2004. Turnover is remarkably high: 8,654 firms enter and 7,896 firms leave between 1978 and 2004. This corresponds to an annual rate of entry and exit of 10.2 percent and 9.2 percent, respectively.

The table also shows the reasons why firms leave the sample. Based on the delisting codes reported in CRSP we classify each exit as either reorganization or default. Default occurs if a firm leaves the sample because of liquidation (delisting codes 400-490), bankruptcy (delisting codes 572 and 574), fraud (delisting codes 580-585 and 587), or non-compliance with the exchange's listing rules (delisting codes 550, 560, 561, and 591). All other exits are classified as reorganizations. These are mostly mergers (delisting codes 200-290) and exchanges for other securities (delisting codes 300-390). Over the

Table 1: Turnover and exit reasons

period	Firms beginning	New entrants	Total exit	Exit reasons	
				Reorg.	Default
1978-1980	2285	377	293	259	34
1981-1985	2316	1289	927	732	195
1986-1990	2681	1511	1262	928	334
1991-1995	2933	2321	1040	684	356
1996-2000	4407	2523	2906	2061	845
2001-2004	3195	624	1468	878	590
End of 2004	2923				
Total	-	8645	7896	5542	2354

This table groups the sample in various subperiods and shows the number of sample firms at the beginning of each period as well as the number of entering and exiting firms. The two columns to the right of the table show the reasons why firms leave the sample. Based on the delisting codes reported on the CRSP tapes we distinguish between two exit reasons: default and reorganizations. Default is assumed if a firm is liquidated (delisting codes 400-490) or dropped from the exchange because of bankruptcy (delisting codes 572 and 574), fraud (delisting codes 580-585 and 587), or non-compliance with the exchange's listing rules (delisting codes 550, 560, 561, and 591). All other exits are classified as reorganizations.

whole sample period, default accounts for roughly 30 percent of all exits.

The high rates of organizational exit are also evident in Table 2 which shows how many of the sample firms make it from one period to another. For example, only 422 of the 2,285 firms present in the 1978 sample survive to the next century. This corresponds to a survival rate of 18.5 percent. Similarly, only 24 percent of the firms that go public between 1986 and 1990 are still present in 2001-2004.

3.2 Firm age

Our main variable of interest is age, defined as one plus the difference between the year under investigation and the year in which the firm is "born". It is difficult to find an economically meaningful definition of when a firm actually

Table 2: Survival of firm

period	**firms by year of entry**						
	1978	1979-1980	1981-1985	1986-1990	1991-1995	1996-2000	2001-2004
1978	2285						
1979-1980	2099	377					
1981-1985	1854	331	1289				
1986-1990	1266	216	961	1511			
1991-1995	925	131	582	1128	2321		
1996-2000	765	105	426	764	1829	2523	
2001-2004	422	53	229	370	817	1398	624

This table shows the survival of firms by year of entry. On the diagonal are the firms that enter the sample in a given subperiod, as reported in column 3 of Table 1. The cells below the diagonal show how many of these firms survived to next periods. For example, 422 of the 2,285 that were present in 1978 survived to the 2001-2004 period.

comes into existence. Consider the case of Microsoft. One could argue that Microsoft was born in 1975, when Bill Gates first used the name "Microsoft" in a letter to Paul Allen. Alternatively, Microsoft's year of birth could be 1977, the year in which a partnership agreement between the two co-founders was signed. Yet another landmark in the history of Microsoft that could serve as the year of birth is 1981, when Microsoft incorporated. Finally, one could also argue that Microsoft—as we know it today—came into existence by going public on March 13, 1986.

In this paper, we compute age as the time elapsed since the date of listing.[1] Since the CRSP tapes go back as far as 1925, the maximum age a firm can have at the beginning of our sample period in 1978 is 54 years, and 80 years at the end of it in 2004.

The practical reason for our definition is data availability—the date of listing (or first appearance on the CRSP tapes) is the only information related to firm age that is provided by COMPUSTAT. However, the definition makes also economic sense. Listing is a defining moment in company life, "a watershed event in the firm's life-cycle", according to Aslan and Kumar (2007). Not surprisingly, listing affects ownership and capital structure, multiplies growth opportunities, increases media exposure, and demands different corporate governance structures (see, for example, Röell (1996) and Loderer and Waelchli (2008)).

Panel A of Table 3 shows that, on average, our sample firms are 14 years old (i.e., they are still independent organizations 14 years after going public); the median is 10. The sample distribution of firm age remains relatively stable over the sample period, except in the late 1990's when the median drops to 7 years, possibly the reflection of the dot-com IPO-wave.

Panel B of Table 3 shows that the median age of the firms that get delisted

[1] More precisely, we approximate a firm's year of birth with the earliest of: (a) the year in which the firm appears on the CRSP tapes; (b) the year in which the firm is included in the COMPUSTAT tapes; and (c) the year for which we find a link between the CRSP and the COMPUSTAT tapes (based on COMPUSTAT data item LINKDT). If, for example, a firm enters the CRSP or COMPUSTAT tapes in 1996, its age is 1 year at the end of 1996 (1+1996-1996) and 5 years at the end of 2000 (1+2000-1996).

Table 3: Firm age

Sample	**Mean**	**Median**	**Max**	**Stdev**	**N**
Panel A: Firm age by sample year					
1978 - 2004	14.14	10.00	80.00	13.79	82845
1978	14.51	9.00	54.00	12.92	2285
1980	15.78	10.00	56.00	13.09	2277
1985	15.47	14.00	61.00	13.76	2589
1990	14.96	10.00	66.00	13.90	2808
1995	12.92	8.00	71.00	13.88	4097
2000	12.27	7.00	76.00	13.41	3523
2004	15.91	11.00	80.00	14.42	2923
Panel B: Firm age by delisting reason					
All delistings	11.78	8.00	80.00	11.91	8146
Reorganization	12.64	9.00	80.00	12.73	5735
Default	9.73	7.00	69.00	9.40	2411

The table shows descriptive statistics for the age of our sample firms.
Panel A reports firm age by sample years.
Panel B is constraint to those firms that do not survive and shows firm age by the two reasons for delisting (reorganizations and defaults, as defined in Table 1.

is 8 years, compared to 10 years in the full sample.[2] When sorting the sample by exit reason, we find that default typically occurs at the age of 7, whereas reorganization happens at a median age of 9. Based on a rank-sum test, this difference is statistically significant with confidence 0.99.

[2]Note that the number of observations reported in Panel B differs from those in Table 1. This is because in Table 3 we also include the 250 firms from the 2004 sample that delist in 2005.

Table 4: Descriptive statistics

variable	**mean**	**median**	**min**	**max**	**stdev**	N
ROA	0.0764	0.0826	-15.1403	9.1905	0.1852	81139
Tobin's Q	1.75	1.27	0.20	173.64	2.18	81915
Capex-to-sales	0.0950	0.0402	-0.0630	13.4163	0.2455	81298
Size	592.63	83.20	0.00	181737	2790.95	81951
Focus	0.8494	1.0000	0.1129	1.0000	0.2357	82845
Leverage	0.2009	0.1586	0.0000	21.0597	0.2158	82835
R&D-to-sales	0.0686	0.0199	0.0000	18.7053	0.1977	42443
Volatility	0.6652	0.5668	0.0000	11.8138	0.4105	79600

This table shows descriptive statistics for the performance measures and the control variables. Variable definitions are in Table 12.

3.3 Performance measures and control variables

We use two measures of firm performance: return on asset (ROA) and Tobin's Q. Whereas ROA reflects the firm's operating performance, Tobin's Q captures market expectations concerning the firm's growth opportunities. ROA is computed as the ratio of the firm's operating income before depreciation divided by its total assets. Tobin's Q is measured as the market value of the firm's assets divided by their book value.

Our control variables include firm size, growth opportunities, specialization, and volatility. Firm size is the market value of assets in millions of 1978 U.S. dollars—the market value of assets is calculated as the sum of the market value of equity and the book value of debt. Growth opportunities are captured by the firm's capital-expenditures-to-sales ratio as well as its R&D-expenditures-to-sales ratio. Firm specialization is a sales-weighted Herfindahl index of a given firm's different segments, as reported on the COMPUSTAT Segment tapes. Volatility is the standard deviation of the firm's stock return. Table 4 shows descriptive statistics. Variable definitions are reported in Table 12 in the appendix.

4 Survival models

4.1 Econometric background

This section analyzes firm survival and its relation to age and performance. Underlying the analysis is a latent continuous time stochastic process $\{X_t | t \in \mathbb{R}_+\}$ with discrete state space $\mathcal{S} = \{s_0, s_1\}$. The firm is alive in state s_0 and delisted in state s_1. Assume the process is initially in state s_0, i.e. $X_0 = s_0$. After some time $T > 0$ the process moves to state s_1 and the firm is delisted. Thus the life time of the firm, T, also called the duration, is defined as the random variable:

$$T = \inf\{t | X_t = s_1\}.$$

Observations consist of realizations of T.

If one wants to consider alternative reasons why the firm gets delisted (i.e., moves out of state s_0), one has to enlarge the state space. With $K > 1$ alternative exit reasons the cardinality of the state space becomes $K + 1 > 2$ and duration is defined as

$$T = \min\{T_k : k = 1, \ldots, K\}$$

where the latent random variable T_k denotes the duration and s_k the associated exit reason.

The objective of duration or survival analysis is to model the distribution of duration conditional on some exogenous covariates which control for observable heterogeneity. Instead of modelling the *distribution function* of T, $F(t) = \mathbf{P}[T \leq t]$, or the *survivor function* $S(t) = 1 - F(t)$, directly, it turns out to be convenient to examine the *hazard function* instead. The hazard function denoted by λ gives the instantaneous probability of leaving state s_0 conditional on survival up to and including t:

$$\lambda(t) = \lim_{h \downarrow 0} \frac{\mathbf{P}[t \leq T < t + h | T \geq t]}{h} = \frac{f(t)}{S(t)} = -\frac{\mathrm{d} \ln S_t}{\mathrm{d}t}$$

where f denotes the density function of T.[3] The distribution of duration

[3]For the sake of exposition we assume that F is differentiable.

may also be characterized by the *integrated hazard function* Λ:

$$\Lambda(t) = \int_0^t \lambda(t)\mathrm{d}t = -\ln S(t)$$

As pointed out above, a popular class of models is given by the *proportional hazard models* (PH models) or *Cox regression models* defined as follows (see Wooldridge (2002), Cameron and Trivedi (2005), and Florens et al. (2007) among others):[4]

$$\lambda(t; x_{i,t}, \upsilon_i) = \upsilon_i \kappa(x_{i,t})\lambda_0(t) \tag{4.1}$$

where the function $\lambda(t; x_{i,t}, \upsilon_i)$ denotes the hazard of leaving the sample conditional on some, possibly time-varying, covariates $x_{i,t}$. The conditional hazard function λ is obtained by multiplying a so-called baseline hazard function λ_0, which is the same for all firms, by a nonnegative function κ, which controls for firm-specific effects. Typically, this function is of the exponential type: $\kappa(x_{i,t}) = \exp(x_{i,t}\beta)$ where β is a fixed coefficient vector. $\upsilon_i > 0$ allows for unobserved heterogeneity.

Hazard rate models possess several advantages. First, they have been widely used to analyze the turnover and the mobility of firms (see Caves (1998) for a survey). Second, they explicitly incorporate the timing of events, which makes them particularly attractive for analyzing the effect of firm age. Third, they are flexible, since they leave the nature of duration dependence (i.e. the baseline hazard) unrestricted. Moreover, the firm-specific effects have a straightforward interpretation.

$$\frac{\partial \log \lambda(t; x_{i,t})}{\partial x_{i,t}} = \frac{\partial \log \kappa(x_{i,t})}{\partial x_{i,t}} = \beta.$$

The specification (4.1) is not the most general feasible specification, however. Firm-specific effects could change with the exposure to risk. In particular, the coefficient vector β could change over time.[5]

[4] The implementation of these models was inspired by Jenkins (2005) and his Stata routines on `http://www.iser.essex.ac.uk/teaching/degree/stephenj/ec968/`.

[5] The importance of this point is emphasized by Bhattacharjee (2005). He proposes to test some empirical implications of the "passive learning" model of Jovanovic (1982) against the "active learning" model of Ericson and Pakes (1995) (see also Pakes and Ericson (1998)). A time-dependent effect of the initial firm size would help discriminate.

The accelerated failure time model (AFT model) is an alternative but not converse model to Cox's proportional hazard model. In this model the exogenous covariates modify the time scale:

$$T = \kappa(x)^{-1} T_0$$

where T_0 is the normal or baseline time scale. Assuming that the covariates are not time-varying, the implied hazard function becomes:

$$\lambda(t; x) = \kappa(x)\lambda_0(t\kappa(x)).$$

The term accelerated failure model stems from the fact that the baseline hazard is accelerated or decelerated whenever $\kappa(x) > 1$ or $\kappa(x) < 1$. The above transformation shows that without additional assumptions the AFT model does not fall into the class of proportional hazard functions. An important special case arises if duration follows an exponential or Weibull distribution. In this case the AFT model is also a PH model.

A nice feature of the AFT model is that it can be transformed into a regression format. Defining $\mu = \mathbb{E} \ln T_0$, we obtain

$$\ln T = \mu - \ln \kappa(x) + \varepsilon$$

where $\varepsilon = \ln T_0 - \mathbb{E} \ln T_0$. With the additional assumptions that $\ln T_0 \sim \mathrm{N}(\mu, \sigma^2)$ and that $\kappa(x) = \exp(-x\beta)$, we obtain a standard linear regression model:

$$\log(t_i) = \mu + x_i\beta + \varepsilon_i \tag{4.2}$$

where t_i and x_i denote the duration of firm i and the observed characteristics of firm i, respectively. Although equation (4.2) has the form of a standard linear regression model with error term ε_i, a Tobit analysis is necessary because the sampling scheme is usually subject to right censoring.

4.2 Some preliminary evidence

Figure 1 presents the empirical cumulative distribution functions (ECDF) of logged durations as estimated by the Kaplan-Meier estimator for alternative

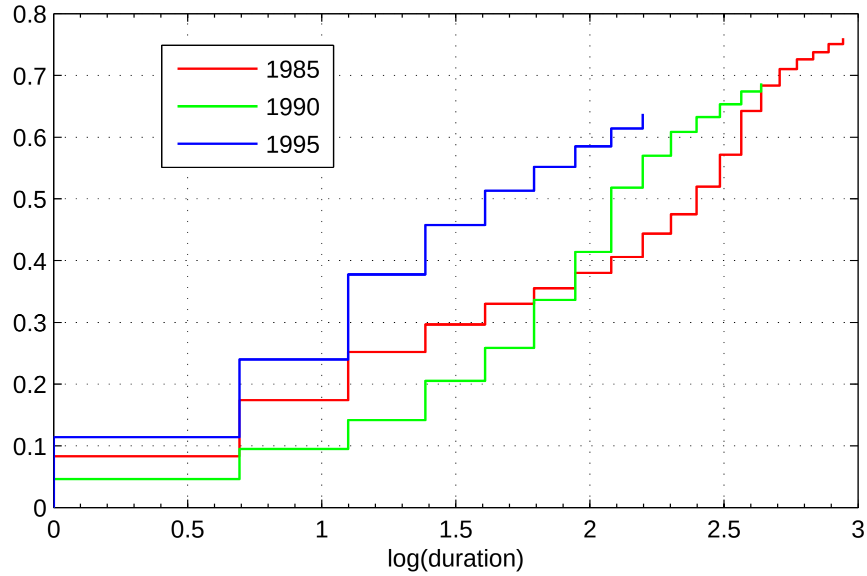


Figure 1: Empirical cumulative distribution functions of logged duration for alternative sampling years

sampling years. This figure shows that the ECDF from the sample with firms existing in 1995 is stochastically dominated by the other two. The probability of observing a duration below a certain value is always higher for this sample.

4.3 Evidence from an Accelerated Failure Time model

The estimates of the AFT model given by equation (4.2) and the specification discussed in Section 2 are reported in Table 5. As there is some ambiguity about the true starting time, we perform the analysis for several sampling periods. The sampling year 1980, for example, covers all firms present on the COMPUSTAT tapes in 1980. To account for industry-specific effects, we include dummy variables for the ten SIC divisions.

The results show a positive dependence on age, but a negative one with respect to Age^2. Both effects are significantly different from zero for most sampling years. Life expectancy first increases with age, but then declines for old firms. As shown at the bottom of the table, the turning point is around year 28 for the sampling year 1980, and steadily greater than that for the subsequent sampling years, up to over 60 for the sampling year 1995.

All control variables have the expected sign. Note, however, that we are investigating logged duration. Hence, the coefficients predicted in the hazard-type environment of Section 2 have to be reversed. Narrower strategic focus, higher financial leverage, and more volatile business environments shorten life expectancy. Larger firms, however, have better chances of survival. Similarly, firms with more growth options—as measured by the R&D-to-sales ratio—live significantly longer. Higher profitability raises the odds of survival, too—the fittest survive. Baker and Kennedy (2002) study the stock-price performance of NYSE and AMEX firms that eventually delist with that of firms that survive. They calculate the returns of firms from 10 years before delisting to their delisting date and show that, on average, "the economic grim reaper kills poorly performing firms." We find consistent results.

We have assumed that all types of exit are caused by the same factors. Exit, however, can mean different things. It can both be a sign of success (an attractive takeover target is taken over in an M&A transaction) or failure (a firm that has to declare bankruptcy). In what follows, we want to find out whether the impact of firm age is different for "happy" and "sad" causes of death. Econometrically this amounts to the estimation of a *competing risk model*. Assuming independent risks, we estimate equation (4.2) conditional on the exit cause. We distinguish between two types of exit: default and reorganization.

Table 6 shows the results when conditioning on default as the cause of exit. Compared to Table 5 the signs of the coefficients remain unchanged. In particular, we still find that life expectancy increases up to a certain age then declines. The negative effect of Age^2, however, is more pronounced in this subsample, indicating that Baker and Kennedy's (2002) "economic grim reaper" strikes old firms harder.

Table 7 conditions on reorganization as the cause of exit. Although a formal test rejects the null hypothesis that the coefficients are the same for

Table 5: Tobit estimates of logged *duration* for different sampling years

covariates	**sampling year**			
	1980	1985	1990	1995
Age	0.014*	0.018***	0.013***	0.011***
	(0.008)	(0.006)	(0.004)	(0.004)
$Age^2/100$	-0.025*	-0.031***	-0.012*	-0.008
	(0.014)	(0.011)	(0.007)	(0.007)
Focus	-0.077	-0.203*	-0.403***	-0.317***
	(0.101)	(0.122)	(0.098)	(0.110)
Leverage	-0.232	-0.663***	-0.319***	-0.347***
	(0.187)	(0.162)	(0.103)	(0.104)
ROA	0.440	1.267***	1.674***	1.775***
	(0.293)	(0.254)	(0.181)	(0.189)
ln(*Size*)	0.110***	0.143***	0.070***	0.052***
	(0.019)	(0.021)	(0.014)	(0.016)
R&D-to-sales	4.134***	2.645***	0.784**	0.436*
	(1.271)	(0.694)	(0.374)	(0.232)
Volatility	-0.831***	-0.286**	-0.361***	-0.206***
	(0.131)	(0.132)	(0.049)	(0.063)
Constant	1.571***	1.234***	2.146***	1.647***
	(0.400)	(0.440)	(0.319)	(0.340)
Observations	2199	2491	2689	3751
Uncensored obs.	1659	1696	1726	2136
R^2	0.026	0.029	0.073	0.032
log-likelihood	-3201.6	-3651.6	-3284.0	-4748.0
Turning point	28.40	28.53	54.01	63.85
in years	(1.48)	(1.22)	(5.54)	(9.52)

The specification includes sectoral dummies. Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Tobit regressions for firms that *default*

covariates	**sampling year**			
	1980	1985	1990	1995
Age	0.030	0.042***	0.034***	0.021***
	(0.019)	(0.011)	(0.010)	(0.006)
$Age^2/100$	-0.037	-0.064***	-0.052***	-0.025**
	(0.032)	(0.020)	(0.017)	(0.011)
Focus	-0.370	-0.703***	-0.601***	-0.307*
	(0.228)	(0.227)	(0.223)	(0.172)
Leverage	-0.416	-1.153***	-0.997***	-0.759***
	(0.406)	(0.273)	(0.220)	(0.156)
ROA	2.683***	2.982***	3.676***	3.136***
	(0.625)	(0.419)	(0.425)	(0.280)
ln(*Size*)	0.190***	0.314***	0.311***	0.127***
	(0.045)	(0.038)	(0.034)	(0.025)
R&D-to-sales	6.427**	6.489***	3.536***	1.488***
	(2.752)	(1.357)	(0.981)	(0.415)
Volatility	-1.799***	-0.889***	-0.585***	-0.651***
	(0.248)	(0.198)	(0.091)	(0.083)
Constant	2.005**	2.066***	2.398***	2.143***
	(0.993)	(0.712)	(0.726)	(0.501)
Observations	701	995	1324	1967
Uncensored obs.	312	419	471	669
R^2	0.126	0.170	0.203	0.167
log-likelihood	-746.5	-1028.3	-1236.3	-1658.0
Turning point	40.83	32.36	32.90	41.42
in years	(4.47)	(1.26)	(1.29)	(2.67)

The specification includes sectoral dummies. Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Tobit regressions for firms that *reorganize*

covariates	**sampling year**			
	1980	1985	1990	1995
Age	0.010	0.018**	0.015***	0.015***
	(0.009)	(0.008)	(0.005)	(0.005)
$Age^2/100$	-0.021	-0.0320**	-0.0116	-0.012
	(0.015)	(0.013)	(0.008)	(0.008)
Focus	-0.117	-0.204	-0.451***	-0.391***
	(0.114)	(0.144)	(0.108)	(0.141)
Leverage	-0.239	-0.619***	-0.319***	-0.346**
	(0.216)	(0.204)	(0.121)	(0.141)
ROA	0.207	1.177***	1.069***	1.530***
	(0.337)	(0.331)	(0.214)	(0.271)
ln(*Size*)	0.133***	0.148***	0.028*	0.077***
	(0.022)	(0.026)	(0.016)	(0.020)
R&D-to-sales	4.417***	2.199***	0.242	0.143
	(1.479)	(0.826)	(0.410)	(0.297)
Volatility	-0.920***	-0.044	-0.116*	0.184*
	(0.163)	(0.183)	(0.068)	(0.103)
Constant	1.505***	0.832	2.244***	1.487***
	(0.465)	(0.567)	(0.365)	(0.464)
Observations	1877	2033	2148	3034
Uncensored obs.	1361	1299	1291	1516
R^2	0.026	0.019	0.030	0.016
log-likelihood	-2750.6	-3001.6	-2518.5	-3839.2
Turning point	24.05	28.18	64.95	61.83
(in years)	(2.24)	(1.39)	(8.09)	(7.46)

The specification includes sectoral dummies. Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the two samples, the results are similar. They show again that firm age has a curvilinear relation with duration. However, the coefficient of Age^2 is significant only in one of the four sampling years.

4.4 Evidence from proportional hazard models

The AFT model concentrates on duration. Cox's regression approach looks at the hazard rate. As before, we estimate the model for different time periods to account for the ambiguity concerning the true starting time. The specification uses the same explanatory variables as the AFT model, but permits them to vary with time. We can therefore use logged firm size in the sampling year ($\ln Size_0$) as an additional regressor to discriminate between active and passive learning models (see Jovanovic (1982), Ericson and Pakes (1995), and Pakes and Ericson (1998)). [6]

The results are reported in Table 8. They basically confirm what we learned so far. Consider, for example, the 1990 sample of firms (third column in Table 8). The coefficients of both Age and Age^2 are significantly different from zero—the linear effect is positive and the squared effect negative. The hazard therefore first declines with age and then increases, the turning point occurring just below age 50. We find similar age effects in the other subsamples.

As before, operationally focused firms are more likely to die. This is also true of highly levered firms and firms with high stock return volatility. Profitability, size, and growth options, however, reduce a firm's propensity to die. Interestingly, the firm's initial endowment, as measured by its size in the sampling year ($\ln(Size_0)$), has a *positive* effect on the hazard function, contrary to both passive and active learning models and to the findings in Bhattacharjee (2005).

We also checked the assumption of proportional hazard by performing the

[6] We also experimented with aggregate economy-wide variables like GDP growth or the interest rate spread between commercial papers and T-bills. As their coefficients turned out to be statistically insignificant, we have not included these variables in the results reported below. Moreover, we do not correct for unobserved heterogeneity.

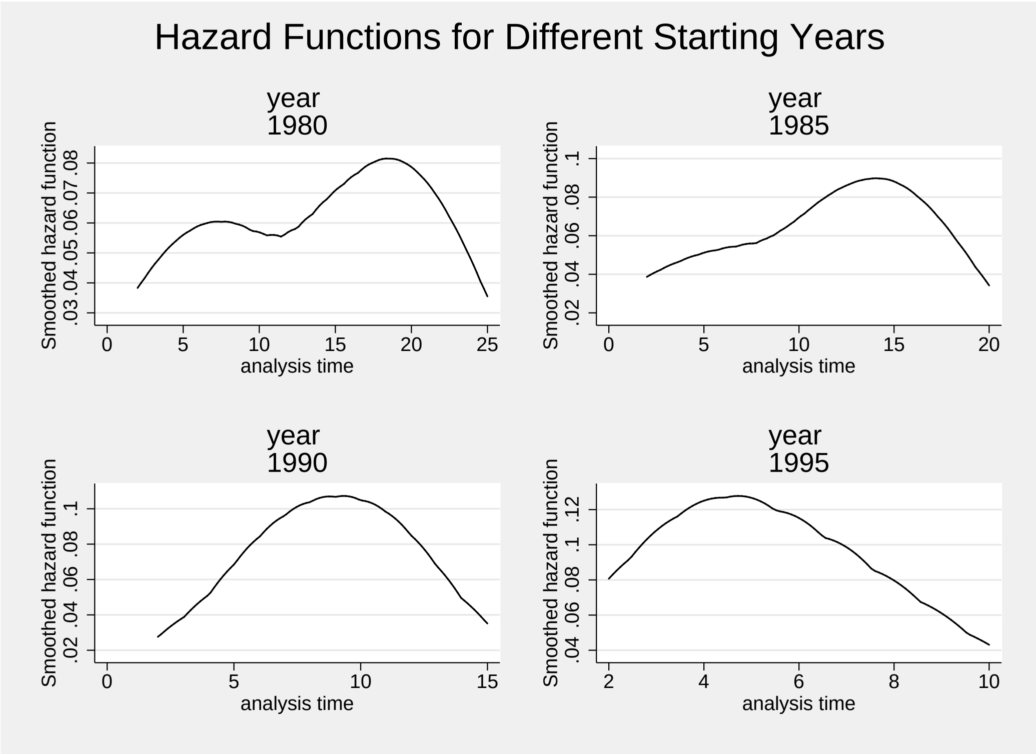


Figure 2: Smoothed baseline hazard function for different starting years corresponding to Table 8

Grambsch-Therneau global test (Grambsch and Therneau (1994)). Under the null hypothesis of proportionality the test statistic is asymptotically χ^2 with 18 degrees of freedom taking the sectoral dummies into account. The five percent critical value is therefore 28.87. The line labelled "Proportionality Test" reports values of the test statistic. The proportionality assumption must be rejected in only one of the four sampling years, namely in 1995.

Finally, Figure 2 shows the estimated smoothed baseline hazard functions. Except for the sampling year 1980, they are unimodal: first increasing and then decreasing. They also suggest that, all else being the same, the exit hazard reaches its peak in 2000, when many hi-tech firms go out of business.

Although our theoretical model is a continuous time model, data are discrete, since we only know that the exit occurs in a certain year. We are

Table 8: Cox's regression for different starting years with time-varying covariates

covariates	sampling year			
	1980	1985	1990	1995
Age	-0.008	-0.016***	-0.015***	-0.009**
	(0.008)	(0.006)	(0.005)	(0.004)
$Age^2/100$	0.011	0.018*	0.015*	0.009
	(0.013)	(0.009)	(0.009)	(0.007)
$Focus$	0.026	0.367***	0.312***	0.209*
	(0.101)	(0.112)	(0.117)	(0.109)
ROA	-1.517***	-1.489***	-1.322***	-1.188***
	(0.245)	(0.203)	(0.201)	(0.158)
$Leverage$	0.549***	0.785***	0.841***	0.667***
	(0.148)	(0.125)	(0.115)	(0.101)
$\ln(Size)$	-0.232***	-0.218***	-0.200***	-0.215***
	(0.030)	(0.030)	(0.028)	(0.031)
$\ln(Size_0)$	0.107***	0.110***	0.102***	0.145***
	(0.031)	(0.032)	(0.028)	(0.031)
R&D-to-sales	-2.439***	-2.112***	-0.693	-0.542**
	(0.912)	(0.630)	(0.439)	(0.260)
$Volatility$	0.266***	0.283***	0.228***	0.253***
	(0.0818)	(0.0633)	(0.043)	(0.062)
Observations	23451	22391	21285	18504
log-likelihood	-11407.3	-12144.0	-12575.6	-17127.6
Wald χ^2 goodness-of-fit	324.0	434.3	411.9	370.1
Proportionality Test	27.54	18.70	21.59	36.19
Turning point	37.40	42.55	48.38	50.01
(in years)	(4.79)	(3.16)	(4.37)	(6.97)

The specification includes sectoral dummies. Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

thus faced with *grouped duration data*. To account for this feature of the data, we formulate a discrete time representation of the continuous proportional hazard model known as the complementary log-log model (cloglog model; see Jenkins (2005) or Cameron and Trivedi (2005, p. 603)). Apart from the discretization, we retain the previous specification. The results are reported in Table 9. They remain basically unaffected. This is true also of the turning points. The only noteworthy result is that the coefficient of $Focus$ changes its sign in the 1980 and 1995 subsamples.

The cloglog model can be extended to account for unobserved heterogeneity, a problem we might face because of omitted variables and/or measurement error. To this end we follow the nonparametric discrete mixture model of Heckman and Singer (1984) and assume that the population of firms is composed of two subpopulations.[7] The probabilities of belonging to one of the two are π_1 and π_2, respectively. Since $\pi_1 + \pi_2 = 1$, there is only one unknown, $\pi = \pi_2$. We parameterize π with $\pi = \frac{\exp(\theta)}{1+\exp(\theta)}$. To avoid over-parameterization, heterogeneity is assumed to affect only the intercept. We therefore have three additional parameters: π or θ, and the intercepts μ_1 and μ_2. For identification purposes, the intercept of the first group, μ_1, is set to zero.

Preliminary results for the nonparametric discrete mixture model are reported in Table 10. They show that unobserved heterogeneity is indeed present, except perhaps for the sampling year 1985. The intercept is always statistically different from zero. Moreover, with the exception of the sampling year 1985, the probability of belonging to the second group is statistically different from zero and one. So that the hypothesis that π_2 falls in the interior of $[0, 1]$ cannot be rejected. As for the coefficients of the covariates, they basically remain unchanged, including the effect of age. Again, the only noteworthy change relates to strategic focus.

[7] We also estimated a model with parametric unobserved heterogeneity where the unobserved component υ_i in equation(4.1) is assumed to be gamma distributed. However, the results are not affected by this choice.

Table 9: Complementary log-log model for different starting years and time-varying covariates

covariates	**sampling year**			
	1980	1985	1990	1995
Age	-0.016**	-0.011*	-0.011*	-0.006
	(0.008)	(0.006)	(0.006)	(0.005)
$Age^2/100$	0.022	0.009	0.008	0.007
	(0.014)	(0.010)	(0.009)	(0.008)
Focus	-0.181*	0.208*	0.109	-0.232**
	(0.105)	(0.126)	(0.127)	(0.111)
Leverage	0.135	0.416***	0.429***	0.643***
	(0.162)	(0.143)	(0.130)	(0.109)
ROA	-1.860***	-1.848***	-1.634***	-0.232***
	(0.227)	(0.198)	(0.199)	(0.037)
$\ln(Size)$	-0.268***	-0.246***	-0.214***	-0.351***
	(0.0312)	(0.031)	(0.029)	(0.028)
$\ln(Size_0)$	0.154***	0.182***	0.166***	0.293***
	(0.033)	(0.034)	(0.030)	(0.030)
R&D-to-sales	-4.166***	-2.212***	-0.997**	0.112
	(0.948)	(0.633)	(0.463)	(0.282)
Volatility	0.547***	0.531***	0.522***	0.496***
	(0.0761)	(0.060)	(0.053)	(0.063)
Observations	18368	17239	16613	14837
log-likelihood	-4712.0	-4511.6	-4477.7	-5244.9
Wald χ^2 goodness-of-fit	8121.1	7606.7	7405.6	6816.3
Turning point	37.16	56.41	67.18	44.63
(in years)	(2.44)	(10.88)	(15.68)	(7.74)

The specification includes sectoral dummies. Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Complementary log-log model with discrete mixture model for unobserved heterogeneity—preliminary

covariates	**sampling year**			
	1980	1985	1990	1995
age	-0.026	-0.030	-0.037	-0.032
	(0.012)	(0.008)	(0.008)	(0.007)
$age^2/100$	0.033	0.034	0.037	0.031
	(0.020)	(0.012)	(0.013)	(0.012)
$Focus$	-0.372	0.211	0.023	-0.199
	(0.140)	(0.144)	(0.162)	(0.151)
Tobin's Q	-0.157	-0.108	-0.086	-0.060
	(0.052)	(0.039)	(0.034)	(0.022)
$leverage$	-0.094	0.421	0.266	0.138
	(0.195)	(0.156)	(0.165)	(0.155)
$\log(sales)$	-0.708	-0.535	-0.647	-0.761
	(0.060)	(0.068)	(0.062)	(0.058)
$\log(sales_0)$	0.444	0.394	0.444	0.595
	(0.064)	(0.060)	(0.061)	(0.062)
observations	18602	17419	16986	15498
log-likelihood	-4852.51	-4683.37	-4677.73	-5638.88
μ_2	2.091	1.195	2.005	2.892
	(0.199)	(0.344)	(0.255)	(0.188)
π_2	0.658	0.716	0.450	0.372
	(0.041)	(0.350)	(0.046)	(0.024)
turning point	38.97	43.72	49.83	50.71
in years	(2.62)	(2.21)	(2.90)	(3.40)

The specification includes sectoral dummies
π_2 is the probability to belong to the second subpopulation
μ_2 is the intercept for the second subpopulation
estimated standard deviations in parenthesis

5 Firm age and performance

We have seen that a firm's exit hazard typically decreases at young age and then increases as firms grow older. We also saw that better performance increases the odds of survival, consistent with the notion that the fittest survive. We do not know, however, how age affects performance. Do firms get better over time, or do they eventually deteriorate?

Table 11 shows the results of robust panel regressions of firm performance on age. Performance is measured alternatively with return on assets (ROA) and Tobin's Q. Both performance metrics are measured as deviations from the median industry value (based on two-digit SIC codes). To allow for a curvilinear relation between age and performance, age enters the regression linearly and with its squared value. We estimate the regression specification discussed in Section 2. We control for firm size, strategic focus, capital expenditures, technological sophistication, and volatility. Firm fixed effects account for the heterogeneity of firms, and period fixed effects capture the impact of the overall state of the economy. Regressions (1) and (2) are estimated for the full sample; regressions (3) and (4) exclude all firms that operate in the financial industry (SIC codes 6000-6999).

The relation between age and performance has the predicted shape. It first declines and then goes back up. Both the linear and the squared effect are statistically significant with confidence 0.99 regardless how we measure performance and whether or not we include financials. The turning point, i.e., the year in which the positive nonlinear effect of age starts overcoming the negative linear effect, however, comes relatively late. In the case of the first regression, which measures performance with ROA, it takes 70 years after the IPO until performance recovers with age again. When we measure performance with Tobins's Q, it takes an estimated 57 years until an additional year of age improves performance.

Figure 3 illustrates the implications from these findings by plotting the relation between firm age and performance in the full sample and ignoring all the other variables in the regression equation. Figure 3(a) refers to the first regression in Table 11 with the ROA-measure of performance. Figure 3(b)

Table 11: Firm age and performance

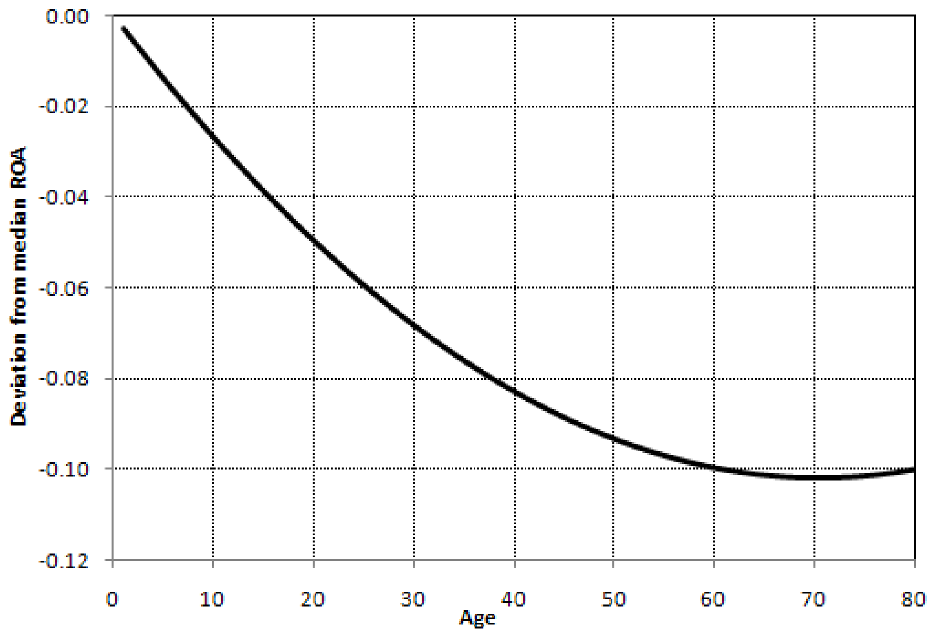
	Full sample		Non-financial firms	
	ROA	Tobin's Q	ROA	Tobin's Q
Age	-0.003***	-0.047***	-0.003***	-0.048***
	(0.0003)	(0.003)	(0.0003)	(0.003)
$Age^2/100$	0.002***	0.041***	0.002***	0.043***
	(0.0004)	(0.005)	(0.0004)	(0.006)
Focus	0.023***	0.297***	0.022***	0.300***
	(0.004)	(0.056)	(0.004)	(0.057)
$\ln(Size)$	0.029***	0.388***	0.030***	0.382***
	(0.002)	(0.024)	(0.002)	(0.025)
$\ln(Size_{t-1})$	-0.026***	-0.320***	-0.026***	-0.314***
	(0.001)	(0.024)	(0.002)	(0.026)
Capex-to-sales	-0.031***	0.033	-0.031***	0.023
	(0.005)	(0.037)	(0.005)	(0.037)
$Capex\text{-}to\text{-}sales_{t-1}$	-0.033***	-0.193***	-0.033***	-0.184***
	(0.009)	(0.039)	(0.009)	(0.040)
R&D-to-sales	-0.187***	-0.809***	-0.185**	-0.613**
	(0.066)	(0.277)	(0.072)	(0.239)
Volatility	-0.065***	0.062	-0.068***	0.063
	(0.007)	(0.064)	(0.007)	(0.074)
Constant	0.043***	0.144	0.046***	0.149
	(0.009)	(0.098)	(0.009)	(0.104)
Observations	55995	57005	51716	52565
F-Test	29.79	18.70	28.40	17.59

Variable definitions are in Table 12. The proxies for firm performance are return on assets (ROA) and Tobin's Q. Both performance measures are adjusted for industry effects by subtracting the performance of the median firm in the industry, computed based on all firms in the same two-digit SIC code with sufficient available information on COMPUSTAT. For those firms that do not report their R&D expenses (approximately 50 percent of the sample firms) we assume a value of zero for R&D-to-sales. The results do not change if we estimate the regressions for those firms that report R&D expenses. We use panel regressions with fixed effects and robust standard errors. The overall state of the economy is captured with period fixed effects (year dummies). Regressions (1) and (2) are estimated for the full sample. Regressions (3) and (4) are constraint to the sample firms that do not operate in the financial industry. Robust standard errors are reported in parentheses. The symbols ***, **, and * indicate statistical significance with confidence 0.99, 0.95, and 0.90, respectively. The years under investigation are 1978 to 2004.

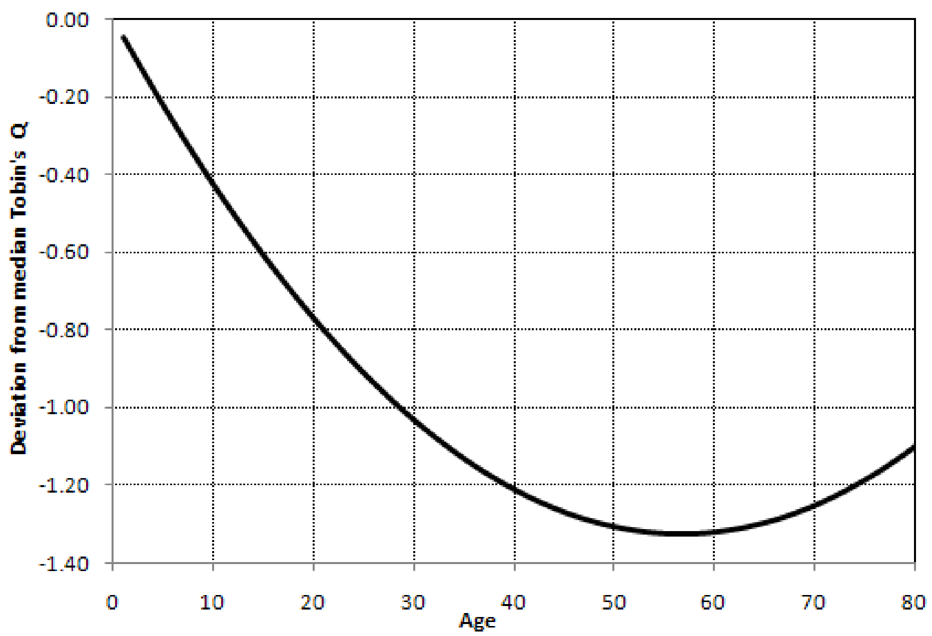
examines performance with Tobin's Q using the coefficients from the second regression. Two aspects are noteworthy. First, compared to what we observe for young firms, the overall contribution of age to performance is negative over the observed range of ages in the sample (the maximum age is 80). Second, only very few firms in the sample actually live long enough to experience the turning point in the curvilinear age-performance relation, since only four out of 100 firms in the sample reach age 50.

The control variables in Table 11 have interesting coefficients that are not always in line with our predictions. Strategic focus has a positive and significant effect on performance in all regressions. Unrelated diversification is therefore bad for business, consistent with the extant literature on the corporate diversification discount (see, among others, Comment and Jarrell (1995), Lang and Stulz (1994), Campa and Kedia (2002), and Villalonga (2004)). Firm size, however, has an ambiguous impact on performance. The lagged value has the predicted negative relation, whereas the current value has a positive relation. We also find that firms with higher current or lagged capital expenditures tend to fare significantly worse, especially when we measure performance with ROA. This is inconsistent at first blush with the claim that corporate investments offset obsolescence. The result could also reflect the fact that higher investments imply higher depreciation charges and therefore lower ROA-measures of performance by definition. The problem with this interpretation is that Tobin's Q correlates negatively with lagged investments as well. Tobin's Q is unrelated to depreciation. Moreover, the results suggest that larger outlays for research and development weaken performance, too. This result is in contrast with previous studies, which generally find that growth opportunities have a positive impact on performance (see, for example, Mehran (1995)). The extant literature, however, focuses mostly on relatively large and established firms (e.g., the S&P500 firms). It could be that smaller firms with large R&D budgets are still in the beginning of their product cycle and therefore still waiting for commercial success. Finally, volatility seems to have a negative impact on performance in our estimates, at least when measuring performance with ROA.

Overall, it seems that performance gradually decreases as firms grow



(a) Age and ROA



(b) Age and Tobins'Q

Figure 3: Relation between age and performance based on regressions (1) and (2) of Table 11

older. This result is robust to the time period we consider. Table 13 shows the coefficients of Age and Age^2 and the associated standard errors if we replicate the regressions from Table 11 for different subsamples and different subperiods, namely 1978-1985, 1986-1995, and 1996-2004. The coefficient of Age is always negative and significant with confidence 0.99, no matter what time period we select. The coefficient of Age^2 is positive and significant with confidence 0.95 or better, except for the period 1996-2004, where it is statistically zero in the ROA regressions. The overall message remains the same: performance declines with age. Based on the coefficients we estimate, it is in principle possible for very old firms to outperform the very young ones, but this would occur at ages of 120 or so, and no firm in the sample has actually lived that long yet.

Loughran and Ritter (1995), among others, document that the stocks of newly listed firms significantly underperform in the years after the IPO (see, however, Fama (1998)). To find out whether our findings are driven by the very young firms, we re-estimate the regressions in Table 11 and alternatively exclude all firms younger than 3, 5, and 10 years. The results remain the same (not shown). In particular, the coefficient of Age is always negative and significant and that of Age^2 is always positive and significant.

6 Conclusion

The purpose of this paper is to investigate how age affects the chances of survival of an organization and its financial performance. Firms learn over time how to do things better and how good they are, but they also tend to lose their flexibility and their initial endowments become obsolete. Hence, survival is partly the result of a tradeoff between learning, new investment, and decay (Agarwal and Gort (1996, 2002)). The results show that survival is a curvilinear function of age. It goes up at younger ages, and then it declines. Older age eventually leads to organizational death. The fittest, however, survive. However, performance weakens with age. For all intents and purposes, survivors are the fittest in a cohort of deteriorating firms.

References

R. Agarwal and M. Gort. Firm and product life cycles and firm survival. *American Economic Review, Papers and Proceedings*, 92:184–190, 2002.

R. Agarwal and M. Gort. The evolution of markets and entry, exit and survival of firms. *Review of Economics and Statistics*, 78:489–498, 1996.

K. J. Arrow. The economic implications of learning by doing. *American Economic Review*, 29:155–173, 1962.

H. Aslan and P. Kumar. Going public and going private: What determines the choice of ownership structure? Working paper, University of Houston, 2007.

G. P. Baker and R. E. Kennedy. Survivorship and the economic grim reaper. *Journal of Law, Economics, and Organization*, 18:324–361, 2002.

E. Bartelsman, S. Scarpetta, and F. Schivardi. Comparative analysis of firm demographics and survival: Evidence from micro-level sources in OECD countries. *Industrial and Corporate Change*, 14:365–391, 2005.

P. Berger and E. Ofek. Diversification's effect on firm value. *Journal of Financial Economics*, 37:39–65, 1995.

A. Bhattacharjee. Models of firm dynamics and the hazard rate of exits: Reconciling theory and evidence using hazard regression models. CRIEFF Discussion Papers 0502, Centre for Research into Industry, Enterprise, Finance and the Firm, 2005.

A. C. Cameron and P. K. Trivedi. *Microeconometrics: Methods and Applications.* Cambridge University Press, New York, 2005.

J. M. Campa and S. Kedia. Explaining the diversification discount. *Journal of Finance*, 57:1731–1762, 2002.

R. E. Caves. Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature*, 36:1947–1982, 1998.

G. L. Clementi. Ipos and the growth of firms. Working paper, Carnegie Mellon University, 2002.

R. Comment and G. Jarrell. Corporate focus and stock returns. *Journal of Financial Economics*, 37:67–87, 1995.

T. Cooley and V. Quadrini. Financial markets and firm dynamics. *The American Economic Review*, 91:1286–1310, 2001.

T. Dunne, M. J. Roberts, and L. Samuelson. The growth and failure of u.s. manufacturing plants. *Quarterly Journal of Economics*, 104:671–698, 1989.

R. Ericson and A. Pakes. Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62:53–82, 1995.

D. S. Evans. The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *Journal of Industrial Economics*, 35: 567–581, 1987.

E. F. Fama. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49:283–306, 1998.

J.-P. Florens, D. Fougère, and M. Mouchart. Duration models and point processes. Discussion Paper No. 2971, Institute for the Study of Labor, Bonn, Germany, 2007.

P. M. Grambsch and T. M. Therneau. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81:515–526, 1994.

J. J. Heckman and B. Singer. A method for minimizing the impact of distributional assumptions in econometric models of duration data. *Econometrica*, 52:271–320, 1984.

B. A. Jain and O. Kini. The post-issue operating performance of ipo firms. *Journal of Finance*, 49:1699–1726, 1994.

S. P. Jenkins. Survival analysis. Lecture Note, 2005.

B. Jovanovic. Selection and the evolution of industry. *Econometrica*, 50: 649–670, 1982.

L. H. P. Lang and R. Stulz. Tobin's q, coporate diversification, and firm performance. *Journal of Political Economy*, 102:1248–1280, 1994.

D. Leonard-Barton. Core capabilities and core rigidities: A paradox in mananging new product development. *Strategic Management Journal*, 13:111–125, 1992.

C. Loderer and U. Waelchli. Protecting minority shareholders: Listed versus unlisted firms. Working paper, Universitaet Bern, 2008.

T. Loughran and J. Ritter. The new issues puzzle. *The Journal of Finance*, 50:23–51, 1995.

H. Mehran. Executive compensation structure, ownership, and firm performance. *Journal of Financial Economics*, 38:163–184, 1995.

W. H. Mikkelson, M. Partch, and K. Shah. Ownership and operating performance of companies that go public. *Journal of Financial Economics*, 44: 281–307, 1997.

A. Pakes and R. Ericson. Empirical implications of alternative models of firm dynamics. *Journal of Economic Theory*, 79:1–45, 1998.

D. Ricardo. *On the Principles of Political Economy and Taxation.* John Murray, London, 1817.

A. Röell. The decision to go public: An overview. *European Economic Review 40*, 40:1071–1081, 1996.

H. Servaes. The value of diversification during the conglomerate merger wave. *Journal of Finance*, 51:1201–1225, 1996.

A. Smith. *The Wealth of Nations.* Modern Library, New York, 1776 [1965].

P. Thompson. Selection and firm survival: Evidence from the shipbuilding industry, 1825-1914. *Journal of Economics and Statistics*, 87:26–36, 2005.

B. Villalonga. Diversification discount or premium? new evidence from the business information tracking series. *Journal of Finance*, 59:475–502, 2004.

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

Table 12: Variable definitions

Name	**Description**
Age	Age is computed as one plus the difference between the year under investigation and the firm's year of birth. The year of birth is computed as the minimum value of: (a) the first year the firm appears on the CRSP tapes; (b) the first year the firm appears on the COMPUSTAT tapes; and (c) the first year for which we find a link between the CRSP and the COMPUSTAT tapes (based on COMPUSTAT data item LINKDT)
Capex-to-sales	The ratio of capital expenditures (DATA178) to net sales (DATA12). The data are from COMPUSTAT
Size	The market value of the firm's assets, calculated as the sum of the market value of equity and the book value of debt, expressed in millions of 1978 U.S. dollars. The data are from COMPUSTAT. We compute the market value of equity as (DATA25×DATA199). The book value of debt is data DATA9. This information is from COMPUSTAT (DATA25×DATA199) and expressed in millions of 1978 U.S. dollars
Focus	The Herfindahl index, H_E, captures the degree of specialization based on the sales of its different segments, as reported on the COMPUSTAT Segment tapes: $H_E = \sum_{i=1}^{N} p_i^2$, where N is the number of different segments, the subscript i identifies the segments, and p_i is the fraction of the firm's total sales in the segment in question
Tobin's Q	Tobin's Q, computed as the market value of the firm's assets divided by their book value. The market value of the assets is approximated by the book value of assets (DATA6) minus the book value of common equity (DATA60) plus the market value of common equity (DATA25×DATA199). The data are from COMPUSTAT
Leverage	The firm's leverage defined as the ratio of long-term debt (DATA9) to total assets (DATA6)
ROA	Return on asset computed as the ratio of the firm's operating income before depreciation (DATA13) divided by total assets (DATA6). The data are from COMPUSTAT. We deflate ROA by subtracting the annual rate of inflation
R&D-to-sales	The ratio of capital expenditures (DATA178 to net sales (DATA12). The data are from COMPUSTAT
Sales	The firm's net sales (DATA12) expressed in millions of 1978 U.S. dollars
Volatility	The annualized volatility of the firm's daily stock return. We calculate the volatility over a one-year window and include all firm-years with at least 100 daily returns. The data are from the daily CRSP tapes.

Table 13: Robustnes check of the relation between firm age and performance

covariates	**full sample**		**non-financial firms**	
	ROA	Tobins's q	ROA	Tobins's q
		sample 1978 - 1985		
Age	-0.004***	-0.036***	-0.003***	-0.035***
	(0.001)	(0.005)	(0.001)	(0.005)
$Age^2/100$	0.006***	0.045***	0.006***	0.0447***
	(0.001)	(0.009)	(0.002)	(0.009)
Turning point	27.95	39.87	24.44	39.34
(Years)	(1.57)	(1.38)	(1.54)	(1.46)
		sample 1986 - 1995		
Age	-0.002***	-0.022***	-0.002***	-0.025***
	(0.001)	(0.005)	(0.001)	(0.005)
$Age^2/100$	0.002**	0.032***	0.002**	0.035***
	(0.001)	(0.008)	(0.001)	(0.008)
Turning point	62.93	35.39	68.22	35.99
(Years)	(5.75)	(1.46)	(7.12)	(1.41)
		sample 1996 - 2004		
Age	-0.005***	-0.068***	-0.006***	-0.0673***
	(0.001)	(0.015)	(0.001)	(0.016)
$Age^2/100$	0.003	0.057**	0.003	0.056**
	(0.002)	(0.023)	(0.002)	(0.026)
Turning point	103.77	59.40	96.26	59.90
(Years)	(16.44)	(4.06)	(14.49)	(4.55)

The table replicates the regressions from Table 11 for different subperiods. To preserve space, we only report the coefficients and associated robust standard errors (in parentheses) for *Age* and Age^2. The subperiods we consider are 1978-1985, 1986-1995, and 1996-2004, respectively. The symbols ***, **, and * indicate statistical significance with confidence 0.99, 0.95, and 0.90, respectively.