ANALYSIS OF SURVIVAL AMONG TOP INDUSTRIAL FIRMS IN TURKEY

by

Coşkun Yağız Özyol

Submitted to the Institute of Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

Sabancı University
July 2018

ANALYSIS OF SURVIVAL AMONG TOP INDUSTRIAL FIRMS IN TURKEY

APPROVED BY:

Doç. Dr. İzak Atiyas ....
(Thesis Supervisor)

Doç. Dr. Ozan Bakış ....

Dr. Öğr. Üyesi Esra Durceylan Kaygusuz ....

DATE OF APPROVAL: JULY 19, 2018

ABSTRACT

ANALYSIS OF SURVIVAL AMONG TOP

INDUSTRIAL FIRMS IN TURKEY

COŞKUN YAĞIZ ÖZYOL

Master's Thesis, July 2018

Thesis Supervisor: Assoc. Prof. İzak Atiyas

Keywords: Survival, networks, exports, productivity, profitability

This paper looks at the survival among the top firms in Turkey and analyses which company specific attributes are correlated with survival. Looking at the company specific attributes such as profitability, number of employees, productivity, exporter status, which industry or network a firm belongs to and where it operates, the paper aimed to get a picture of the top thousand firms in Turkey over a period of 35 years. We found that having a higher number of employees, higher productivity, and higher profitability (with one exception), operating out of a major industrial center, as well as belonging to a secular network were all correlated in a statistically significant way with continued survival among the top firms.

ÖZET

TÜRKİYE'DEKİ EN BÜYÜK

SANAYİ FİRMALARINDA SAĞKALIM ANALİZİ

COŞKUN YAĞIZ ÖZYOL

Yüksek Lisans Tezi, Temmuz 2018

Tez Danışmanı: Doç. Dr. İzak Atiyas

Anahtar kelimeler: Sağkalım, ağ, ihracat, verimlilik, karlılık

Bu makalede Türkiye'deki en büyük sanayi firmalarının İSO listesindeki sağkalımını incelendi ve hangi firma özelliklerinin sağkalıma olumlu ve olumsuz etkisi olduğuna bakıldı. Karlılık, çalışanların sayısı, verimlilik, ihracatçılık durumu, firmanın hangi iş ağına bağlı olduğu, hangi sanayi odasına bağlı olduğu gibi firma özelliklerini incelerek, Türkiye'deki en büyük bin firmayı takip ettik. Çalışan sayısının yüksek olmasının, verimlilik ve karlılığın daha yüksek olmasının, sanayi merkezlerindeki odalara bağlı olmalarının, "laik" sanayi ağlarına bağlı olmalarının bir firmanın sağkalımı ile istatiktiksel olarak anlamlı ve olumlu bir ilişkisi olduğunu bulduk.

ACKNOWLEDGEMENTS

I first wish to express my gratitude to my supervisor Assoc. Prof. İzak Atiyas for his guidance and his many and thorough comments, notes and advice as well as his encouragement.

I would also like to thank my jury members Assoc. Prof. Ozan Bakış and Asst. Prof. Esra Durceylan Kaygusuz for their time as well their many insightful comments which made it into the final version of this paper.

It can go without saying that I wish to thank my family for their moral and material support throughout this process.

TABLE OF CONTENTS

TABLE OF FIGURES

1. INTRODUCTION

Survival of firms within various contexts is a well-researched area. The effects of many metrics such as innovation (Buddelmeyer et al. 2016), organizational structure (Audretsch, 1991) or export (Esteve-Perez et al.) have been thoroughly analyzed. In most of these contexts "survival" was meant to describe the continued operation of a firm generally, or the survival of the firm in a certain market. In this paper we look at a different context for survival.

A look at the top firms in any market are a good metric for understanding both the underlying mechanisms, and the changes that economy faces over time, which is why it is very fortuitous to be presented with a consistent set of the top manufacturing firms in Turkey since 1980 compiled by the *Istanbul Sanayi Odasi (*henceforth ISO), which list the top 500 and top 1000 firms (the dataset is further elaborated in the data section) and organized by sales but including many other micro-data, from the number of employees and profits to exporter status. One interesting new area of research this paper has explored has been built on the foundation laid in Atiyas et al. (2016) who used the 2013 membership lists for the Turkish Industry and Business Association (TUSIAD), The Young Businessmen Association of Turkey (TUGIK), Anatolian Businessmen Association (ASKON), Independent Industrialists and Businessmen Association (MUSIAD), Young Businessmen Association of Turkey (TUGIAD), All Industrialists' and Businessmen's Association (TUMSIAD), Turkish Enterprise and Business Confederation (TURKONFED), and Turkish Confederation of Businessmen and Industrialists (TUSKON) to add membership information to each of the firms in the ISO lists for the various business associations across Turkey. Using this data and their previous analysis, we have also separated the firms into belonging to secular and religious networks (more information on this data will be presented in the data section) and looked at the effects of their membership status on their survival among the top firms in Turkey. Further building on this is the matching of the firms on this list with membership in various business associations of both secular and religious affiliations, carried out for Atiyas et al. (2016) which has given us more data for analysis.

Looking at this data allows us to ask what allows a firm to stay at the top in Turkey. As the list is compiled year by year, with many of the firms exiting, and many others staying, we can gage the effects of the many economic indicators with survival among top firms.

Previous research has challenged many of the received ideas about the operational survival of firms such as proposing that innovation has a negative relation with survival (Buddelmeyer et al. 2016). For our research, we propose to test the following hypotheses that higher survival rates are correlated with:

1) Higher productivity and profitability of firms,
2) Being an exporter,
3) Higher number of employees,
4) Belonging to a secular network,
5) Operating out of a major industrial center,
6) Belonging to mid-technology industries such as chemicals, metals, and machinery.

For our hypotheses, (1) and (5), our main basis is simply an intuition that better performing firms, and firms in more competitive environments would be more likely to survive. For our hypothesis (2), we refer to the long literature on “learning-by-exporting”, which tests the assumed positive relation between exporting and productivity. Love et al. (2013), Yasar et al. (2013), Salomon and Shaver (2005) all find evidence (although subtle and ambiguous in some cases) of a positive causal relation between exporting and an increase in productivity. For hypothesis (3), we refer to the conclusion of Lobos et al. (2012) for Poland who found that higher the number of employees, the higher the probability of survival (although again to reiterate, in their case, ‘survival’ means continued operation). For our hypotheses (4) and (6) we refer to the conclusion of Atiyas et al. (2016) who found a productivity gap between firms operating in secular and religious networks, with secular network firms being more productive, and belonging to mid-technology industries. Further for hypothesis (6), we can also suggest that low technology firms require less knowledge capital and have larger demand shocks, and given the economic crises in Turkey, would be more susceptible to the competition from other emerging economies than mid-technology firms: a shift in demand for textiles is much more likely than that for automobiles.

Therefore, this paper will test the hypotheses above using survival analysis. The methodology in this paper will use both the traditional tools of survival analysis with the various time-invariant metrics such as `entry` values for the number of employees, profitability, etc. to test whether these values are related to firm survival in the top, and a less-used time dependent covariates with discrete time model to test the overall effect of the changes in these covariates over time.

2. LITERATURE REVIEW

Theoretically, this paper makes use of the survival analysis models developed first in Kaplan and Meier (1958) and Cox (1972). The first of these was used to estimate the survival of an individual over a period of time non-parametrically – not having assumed a form for the shape of the survival rates of the individuals, and second assuming various exponential forms. Various summations on the shape that the hazard function can take have been expanded upon by Kiefer (1988) and Rodrgiuez (2010). The specific types of distributions that the hazard function can take, and the implications arising from it will be discussed further in the methodology section. This paper partially uses the time-invariant covariate model for survival analysis, where each firm is assumed to have certain characteristics that do not vary with time. Most of the methods that are used to this end are explained by Kiefer and Rodriguez. However, the data we have also can be analyzed as discrete time data, where certain firms have characteristics which change discretely over time (for example, the number of employees of a firm changes over the period it is included in the ISO list), and multivariate regression on this data, while less broadly used, has been first described by Prentice and Gloecker (1978) and then by Jenkins (2004), and are explained in more detail in the methodology section.

Both the Cox hazard model and the Kaplan-Meier non-parametric model have been used for a variety of applications across various fields: survival of heart transplant patients (Crowley and Hu (1972)), survival of heavy machinery (Madeira, Infante, Didelet (2013)), and even survival of animals in the wild (Pollock, Winterstein, Bunck (1989)).

In economics, this method has been used with unemployment duration of individuals: given various covariates such as age, gender, race, education, how long does an individual "survive" in unemployment. These studies either focus on non-Parametric estimations of the survival function in examples such as Ciuca and Matei (2010) or semi-parametric Cox hazard functions, such as Kavkler et al. (2009) or Hoffman (1991).

When it comes to survival of firms, there also is a rich literature. Survival, of course, can have various interpretations on the firm level. A common use of the analysis has been to look at the operational survival of a firm given various covariates. Lobos and Szewczyk (2012) look at the survival of firms in Poland using the Kaplan-Meier analysis to see whether specific groups - such as a firm being run with partners or being in a competitive market among other groups – influence firm survival, here defined as continuing to operate.

The effect of innovation on survival has also been a popular area of study. One such example is Buddelmeyer et al. (2006), using survival analysis, finds that there is a negative correlation between innovation and firm survival. This is in keeping with the findings of Audretsch and Mahmood (2001) who have found that for large firms, a highly innovative environment is negatively correlated with survival and with small firms, there is no stastistically significant correlation. In a different paper, Audretsch (1991) also concluded that technological regime does not have an impact on short-term survival rates. Another example is Cefis and Marsili (2005) look at the effects of innovation on survival. Contrary to this, looking at Spanish data, Esteve-Perez et al. (2010) have concluded that participation in R&D activities increase the chances of a firm's survival. Esteve-Perez. et al. (2010) is also one of the few papers in that employ a complementary log-log analysis to conduct a discrete time analysis on time-varying covariates.

A well-trodden area of research has been the survival among new firms. Some examples of the research done on the subject include Agarwal and Gort (2002) who have looked at the hazard rates of firms survival of different phases of product and firm cycles and have found that firm survival is dependent on product and firm cycles. Bruderl, et al. (1992) has used both non-parametric life tables and a modified log-logistic model to look at the survival rates among newly formed firms, looking both at the human capital and organizational aspects of a business and found that smallness in both capital and labor are a disadvantage to the survival of a firm. Their life tables have also shown a step-function where firm mortality increases in the first few years of business before falling off. Holmes et. al. (2011), also looked at survival among newly established micro-firms and SMEs in the United Kingdom and found that increased plant sizes are positively correlated with firm survival for SMEs but negatively impact micro-firms. Unlike Bruderl, et al., they also concluded that survival initially has a positive duration dependence followed by a negative duration dependence. As a note, we do not expect to find this in our research, as they look at operational survival of new firms, whereas we look at the survival of already established firms among other top firms. Esteve-Perez et. al. (2010) have used parametric and semi-parametric models to look at Spanish data for firm survival and have found that youngest and oldest firms are at the highest risk of failure, and size is positively related to survival.

Another definition of survival is a survival in a certain market. Esteve-Perez et al. (2007) looks at the survival of Spanish firms as exporters given a set of covariates (i.e. the determinants of persistence) and finds that remaining an exporter increases with the spell of exporting (negative duration dependence of the exit event) and firms which export to closer countries have higher survival. Outside of survival in and of itself, exporting has also been investigated in the context of other metrics of firm survival. For instance, a common area of research is the testing the "learning-by-exporting" hypothesis, where the assumption that a firm's productivity increases by being exposed to foreign markets with an increase of knowledge and expertise. Yasar et al. (2007), look at the relationship between exports and productivity in the Turkish Apparel and Motor industries, and find a causality from productivity to exports and a small causality from exports to productivity. Similarly ambiguous findings are present in Love et. al. (2013), who using data from high-tech SMEs in the United Kingdom, conclude that while there is evidence on the effects of learning by exporting, this effect is "subtle and dependent on the export exit and entry behavior".

Some studies in the turnover among top businesses in a country have been done in the context of Fortune 500 in the USA. Some such studies are Shanklin (1986) and Stangler and Arbesman (2012), the latter of whom found that the share of older and larger companies in the Fortune 500 has been growing since 1958.

The Turkish context among the top firms in Turkey has been analyzed and even this particular dataset has been previously used in Atiyas et al. (2016) who have also noted the emergence of new growth centers in Anatolia during the rule of the Justice and Development Party (AKP) and found a productivity gaps between these new centers and the traditional production centers (albeit one which decreased between 2006-2010). Similar results in the productivity gap were also noted by Atiyas et al. which also found that larger firms are more productive than smaller firms across Turkey. These will be kept in mind while looking at survival among the top firms in Turkey across years. Especially the performance of religious networks vis-à-vis secular ones throughout the 2000s will be an area of interest.

3. DATA

The data used in this paper comes from the yearly list of the top firms in Turkey compiled by Istanbul Chamber of Industry (henceforth called ISO). There are two lists made by ISO called the ISO-500 and the ISO-1000, which record the top 500 and 1000 firms in Turkey respectively. The 500-list was compiled between 1980 and 1997. After 1997, the number of firms was expanded to 1000 to produce the 1000-list.

Apart from the names of the firms and their relative positions on the lists, there are some useful microdata that ISO also reports. Total sales, profits from sales, both domestic and international capital (for some years only), loss/profit for the year, the number of employees for the year, the NACE and ISIC industry codes, and the urban chamber of which the firms are members (signifying the headquarter city for the firm), among others are reported.

The financial data for each firm is given in Turkish Liras for each year. To adjust for inflation, we have used another list, compiled by the Turkish Statistical Institute (TURKSTAT) and published by the Central Bank for the Producer Price Indices (henceforth called UFE). These indices are recorded on a monthly basis from 1982 to 2014, taking the 1981 Lira as their basis. For each month, the list gives a 12-monthly percentage change (compared to the same time the previous year). We take the average of all these percentage changes and find one UFE value for the year. By compounding over the years from 1981 to 2014, we produce indices that will give a multiplier to convert all values to 1981 Liras. The prices are adjusted thus:

$$RealPrice = \frac{Price}{UFE_t}$$

Where t is the year for which the price is taken.

Given the amount of raw data from the ISO lists, we also produce some useful covariates for our analysis. First of these is the labor productivity for each firm. It is calculated as:

$$LaborProductivity = \frac{Real\,Value\;Added}{\#\,Employees}$$

The real value added is simply the reported value added of a firm in the list adjusted for inflation. The number of employees for the firms is also reported in the ISO list.

The second covariate we produce is the entry profitability a firm. It has been calculated in two specifications but only the first has been used in the results section. The analysis in both specifications have yielded similar results.

$$Profitability_1 = \frac{RealProfit}{RealRevenue}$$

$$Profitability_2 = \frac{RealProfit}{RealTotalAssets}$$

Here real profit is the reported profit of a firm adjusted for inflation, the real revenue is the yearly reported revenue adjusted for inflation, and the real total assets are the reported total assets adjusted for inflation.

As stated in the introduction, we have also made use of the 2013 membership list of TUSIAD, in order to add membership affiliations for various business networks following Atiyas et. al (2016). There is more detailed reasoning on why various networks were identified with various political ideologies on that paper, so we will only present the specification of religious and secular membership in Table 1:

Table 1: Membership specifications for the firms depending on the various membership statuses

Network	**Multiple Memberships**			
	None	+MUSIAD	+TURKONFED	+TUSKON
ASKON				
MUSIAD				
TUGIAD				
TUGIK				
TUMSIAD				
TURKONFED				
TUSIAD				
TUSKON				
MEMBERSHIP SPECIFICATION				
	Religious Network			
	Secular Network			

4. DESCRIPTIVE STATISTICS

The IS0-500 data has 2093 firms on its list between 1980 and 2014, of which 1672 fail before 2014, leading to an 80% failure rate. The average time at failure for a firm in ISO-500 is 6.94 years.

The IS0-1000 data has 2572 subject firms on its list between 1997 and 2014, of which 1727 fail before 2014, leading to a 67% failure rate. The average time at failure for a firm in ISO-1000 is 6.25 years.

4.1. ISO-500

We first look at the retention rate of firms in the list at cohorts chosen at 5-yearly intervals: what percentage of the firms in a cohort list remains each year until 2014.

Figure 1 shows the retention rate of firms across different yearly lists. For example, the blue line which shows the 1980 cohort has a value around 0.45 at 5 years. This means that, of the firms in the list in 1980, 45% were remaining in 1985. Differently colored lines are the different entry cohorts. Figure 1 shows us that the survival rates of the firms across different cohorts start off widely spread and converge over time that firms remain in the list. We should mention that the 1980 list is an outlier, most probably since 1980 is the first year that the list was compiled and had some irregularities and `kinks` in data collection that would be worked out in the next years. However, even excluding the 1980 cohort, the 1-year retention rates vary between 77% (for the 1985 list) and 89% (for the 2010 list) and show a generally increasing trend across years (see Figure 2). Comparing this to the 15-year retention rates, which vary between 29% and 33%, or 25-year retention rates, which vary between 17% and 19%, shows that cohorts show less variance as firm survival increases. This may be because as time spent in the list increases, the survival rates of the firms tend to be similar, independent of the cohort year.

Another way of visualizing the data is in Figure 2 which shows the rate of firms remaining in the list across different cohorts, and each line is a different survival duration. For example, the blue line shows the 1-year survival rate of firms for each cohort between 1980 to 2013 (i.e. how many of the firms in the cohort year X were in the list in the year X+1). The orange line shows the 5-year survival rate, etc.

As mentioned before, the 1-year survival rates tend to increase over the years. For example, between the first year when the list was compiled and the year after (between 1980-1981), the drop-off rate was 34%: of those that are in the list in 1980, 34% have dropped off by 1981. The 1-year survival rate peaks with only 9% of the firms in the 2010 list dropping off in 2011. Even if we assume that the first year of the list to retain some artefacts and issues of data collecting, the drop-off rates are still similar in the following years (29% between 1981-1982, 27% between 1982-1983, etc), and decreasing to the 20%s by mid-1980s and hitting the all-time low in 2011.

The 5-year survival rates also show a similar increase. Of the firms in the list in 1980, 55% have dropped off by 1985. The highest survival happens for the 2009 cohort: only 33% of the firms in 2009 drop off in 2014. For the years after 2009, the 5-year survival is impossible to calculate to see whether the trend increases, as the last year of data used in this paper is from 2014.

Contrary to the 1-year and the 5-year survival rates, we can observe that the 10, 15, 20, 25 and 30-Year survivals remain mostly unchanged around their means. The 10-year survival rate of a firm remains around 42% across the years, 15-year survival rate remains around 31%, 20-year around 23% and 25-year around 18%.

Figure 1: The rate of firms remaining in the list across every year for 5-yearly cohorts in ISO-500

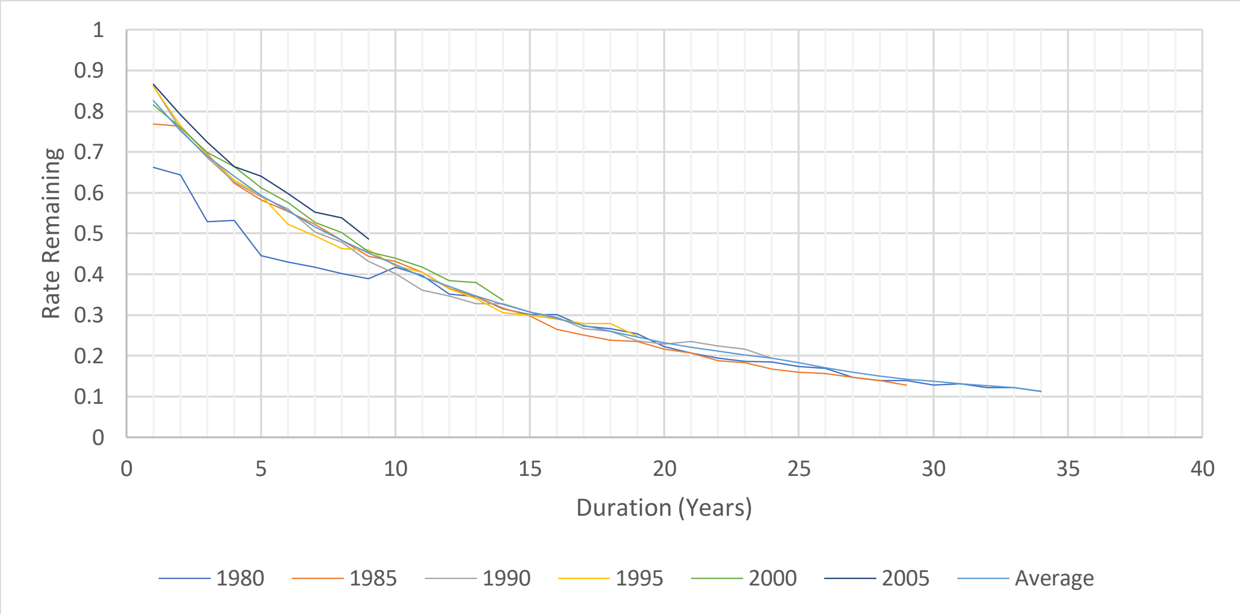
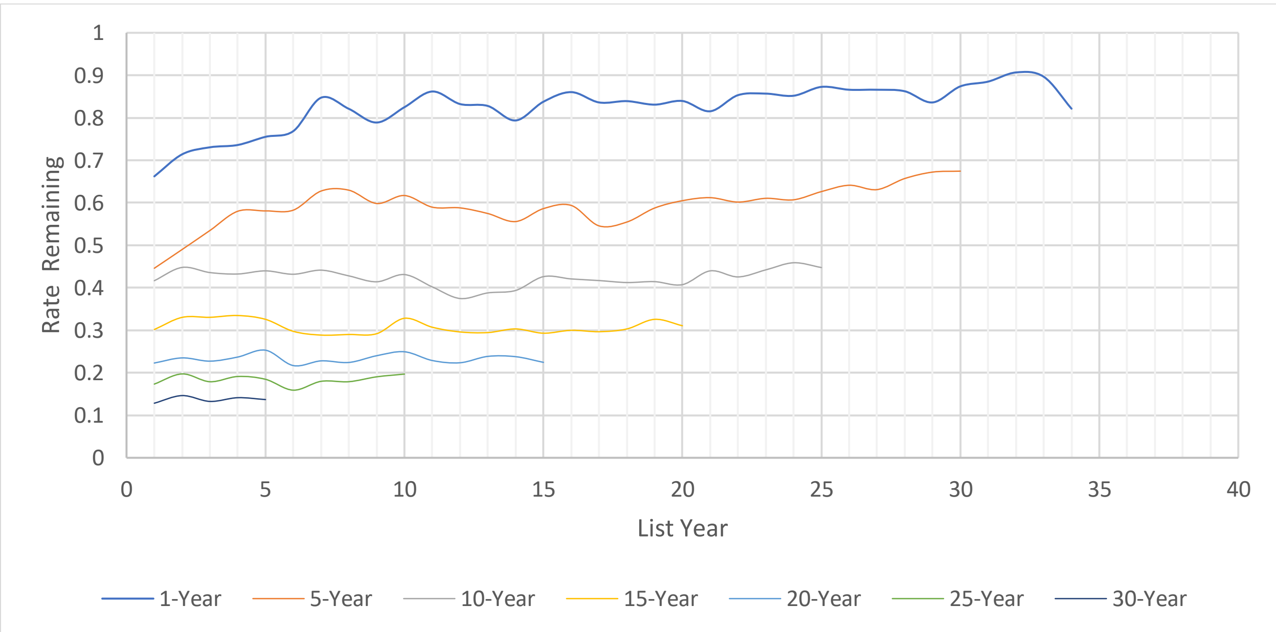


Figure 2: The rate of firms remaining in the list across every cohort for 5-yearly survival durations in IS0-500



4.2. ISO-1000

IS0-1000 list has been compiled since 1997, and therefore, there are fewer years to look through. There are similar trends among the top 1000 firms as there were among the top 500, however. Figure 3 and Figure 4 show this trend in greater detail.

Figure *3* shows that as with the ISO-500 list (shown in Figure 1) across cohorts, the yearly survival rates tend to converge as survival duration increases. The shorter duration that firms remain in the ISO-1000 list are the most volatile. For example, there is a 11% difference across cohorts for 1-year survival and 16% difference across cohorts for 5-year survival, compared to 7% difference for 10-year survival and 2% difference for 15 year survival.

Figure 4 is similar to Figure 2 but done for the ISO-1000 list. At first look, it appears that unlike Figure 2, there is no clear increase in the 1-year rates of firms in the IS0-1000 list. If we take out the 2013 data, however, there is an upward trend. This trend also occurs in the 5-year and 10-year survival rates.

For instance, of the firms that are on the list in 1997, 28% have dropped off by 1998. The 1-Year survival for the larger list also peaks in 2011, with only a 7% of the firms from the previous year dropping off. More drastically, the 5-year drop-off rates decrease from 45% between 1997 and 2002 to 29% between 2008 and 2013, with the 2008 list being the peak for the 5-year survival.

Across both the ISO-500 and IS0-1000 lists, the short duration survival of firms increase across years, but the long duration survival appears to remain the same. This may imply that for short term survival of the firms in the list, since 1980, the ability of the firms to remain in the list has increased but as survival duration increases, factors which account for firm survival do not vary much across years. There is one major caveat to this observation, however. There are fewer observations as the survival years increase since, the data is collected between 1980 and 2014, we have only 5 cohorts for the 30-yearly survival rates of the firms, and those that have been in the list between 1980 and 1984. Therefore, there may be larger trends that would become more apparent as more data is collected. In our parametric estimation of survival in the *Results* section, we will look at whether given controls, how the hazards to survival change with duration further.

Figure 3: The rate of firms remaining in the list across every year for 5-yearly cohorts in ISO-1000

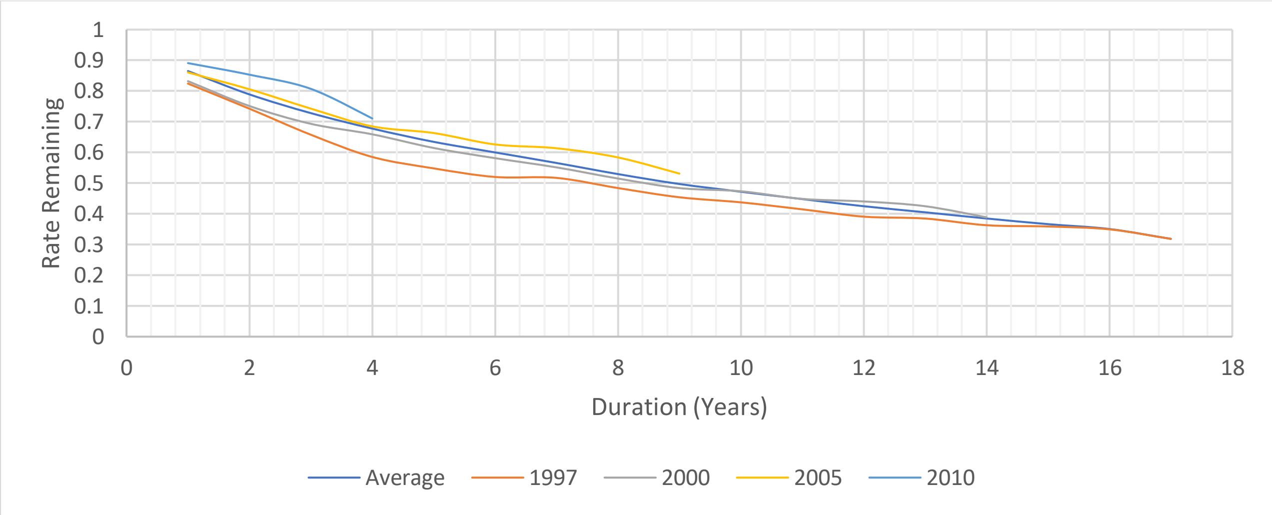
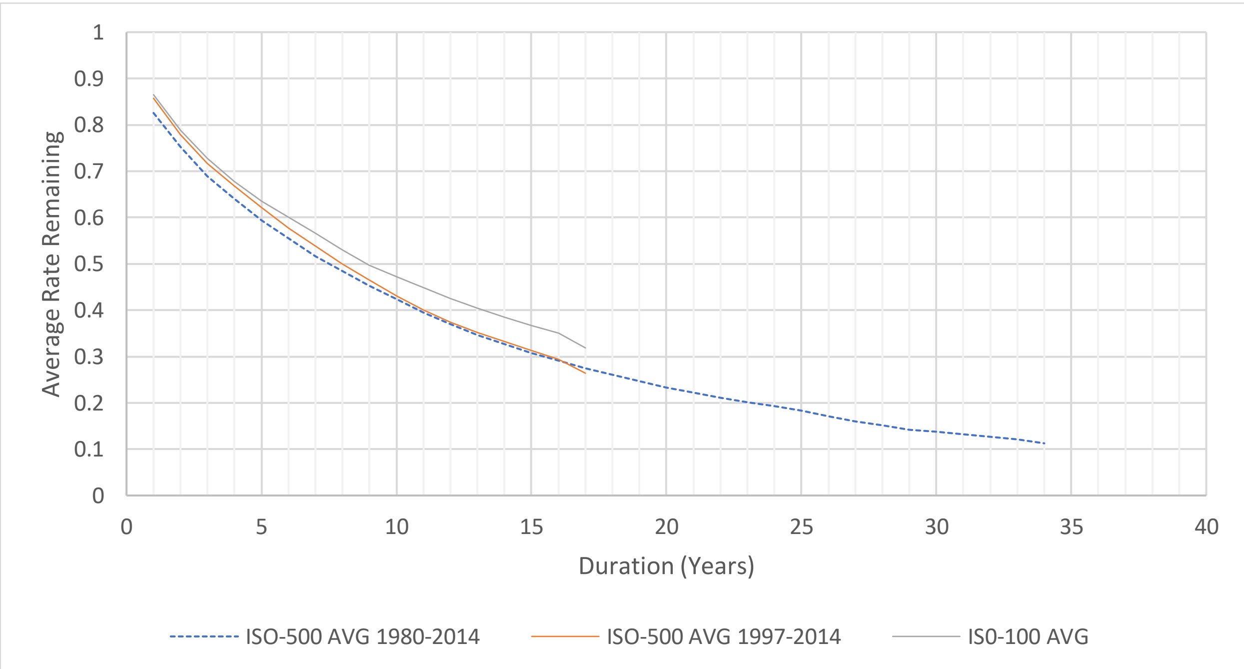


Figure 4: The rate of firms remaining in the list across every cohort for 5-yearly survival durations in IS0-1000

We have also looked at the average yearly survival rates of the firms for the ISO-500 and ISO-1000 lists. In Figure 5, the x-axis shows the survival duration in years, and the y-axis shows the average number of firms remaining for that duration across cohorts. For instance, the 1-year survival rate of firms in 1980 is around 67%, in 1981 it is 64%, etc. We take the average of these rates between the 1980 and 2014 cohorts in the ISO-500 list which gives 83%, which is first point on the dashed blue line.

However, comparing the average survival of all firms in ISO-500 and ISO-1000 may have some issues. For one, the 1-year survival rate average for the ISO-500 list considers all cohorts between 1980 and 2014, whereas the ISO-1000 list considers only the cohorts between 1997 and 2014. Since our previous analysis has shown that the 1-year survival rates have increased between the 1980 cohort and the 1997 cohort, a better method of comparing the two lists may be to simply compare their averages for between 1997 and 2014 – years both lists have existed. In Figure 5, the orange line shows the average survival rates across firms for the ISO-500 starting in 1997, and the grey line shows that for the ISO-1000 list.

Figure 5: Average rate remaining of the firms for different survival durations for the ISO-500 and ISO-1000 lists



We can see that the firms of the ISO-1000 list has a higher average survival rate than the ISO-500 list. Therefore, there is more motion out of the top 500 firms in Turkey than there is out of the top 1000.

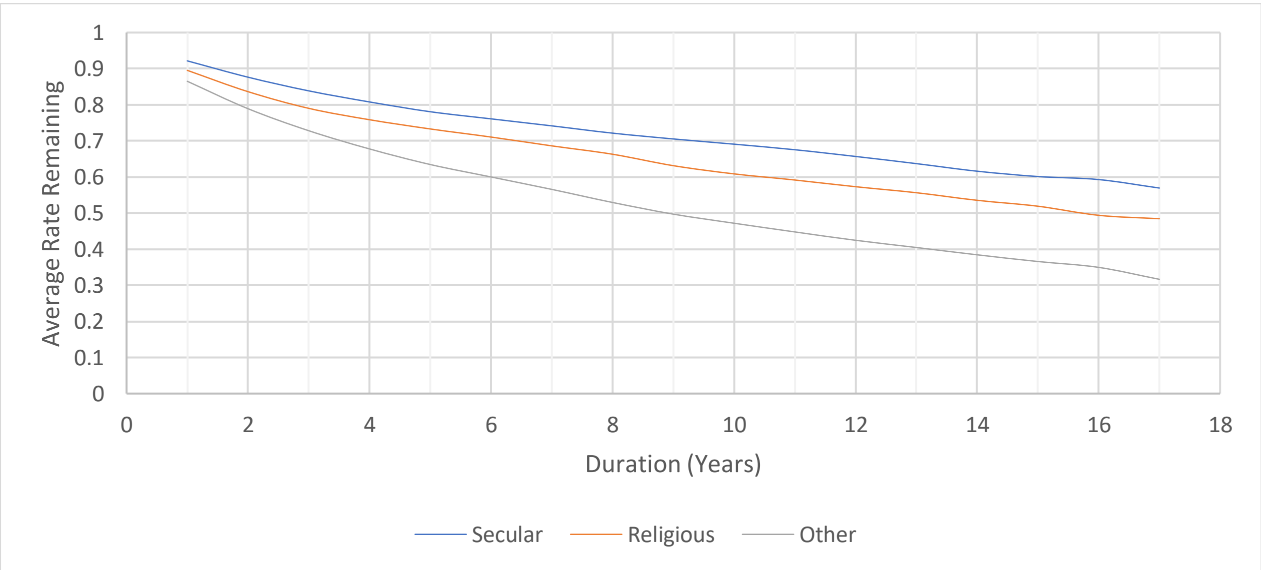
4.3. Network Affiliation

We have also compiled in our list the affiliation of each firm with different types of business networks. To see the survival rates of firms across different networks, firms are given three values for network membership: religious, secular and none depending on the religious affiliation of the business network each belong to. The results for the comparison are shown in Figure 6. We again only show the average rate remaining for different survival durations between 1997 and 2014.

In ISO-1000 between 1997 and 2014, firms in secular networks have performed consistently better than both religious and unaffiliated firms, and firms in religious networks have consistently performed better than unaffiliated firms. Between 1997 and 1998, 18% of all the firms failed to remain in the list. However, this rate is reduced to 12% for firms in religious networks, and 6% for firms in secular networks.

Between 1997 and 2006, the performance of the firms in secular and religious networks converged. For the 2006 cohort, for short term survival duration (less than 5-years), the religious networks perform similarly or better than secular networks. However, in the following years, secular networks are solidly on top again.

Figure 6: The average rate remaining across survival duration by network affiliation in the ISO-1000 list



4.4. Structural Change

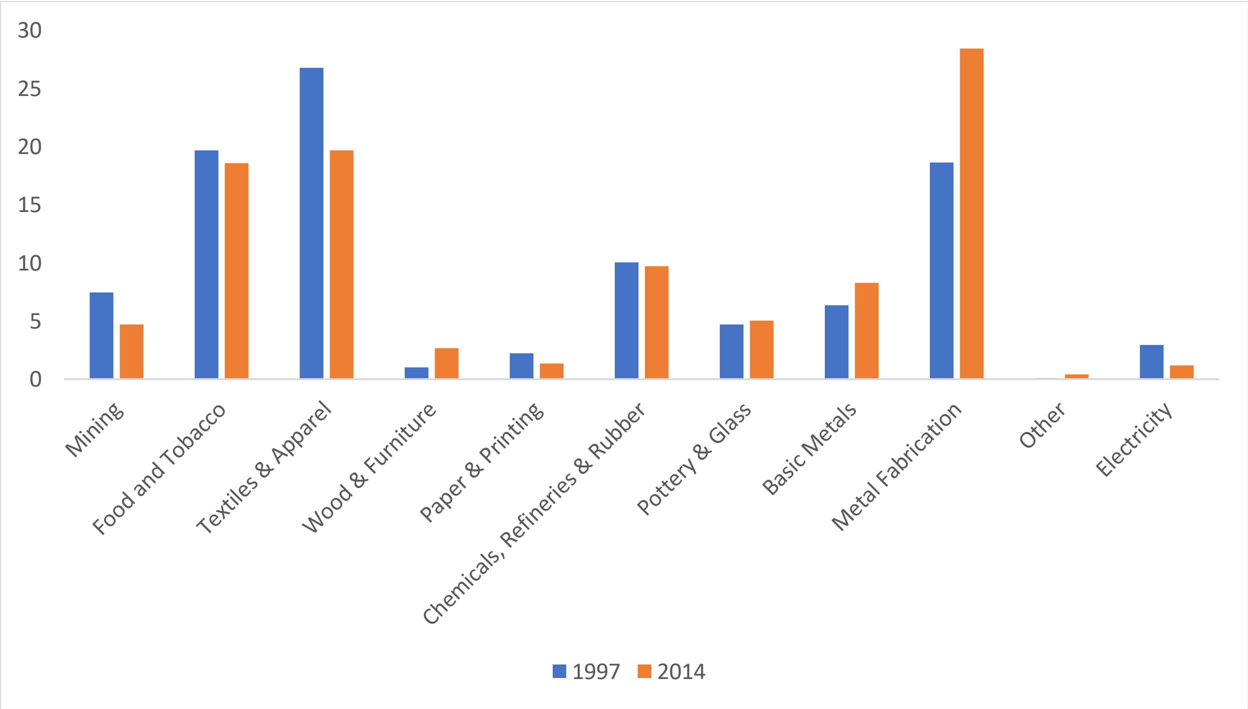
We next look at the structural change in the makeup of the top 500 firms in Turkey. Here, International Classification of All Economic Activities (ISIC) codes have been compiled for the firms in the top 1000 since 1997. The relevant three-digit ISIC codes for the lists between 1993-2014 are listed in the appendix. For ease, we have shortened the ISIC Codes to 2 digits, and have compared the number of employees in the ISO-1000 list between 1997 and 2014. The results are displayed in Table 2:

Table 2: The number of employees by industry in 1997 and 2014

		1997				2014			
ISIC Code	**Desc.**	**Freq.**	**%**	**#Employees**	**% Employees**	**Freq.**	**%**	**#Employees**	**% Employees**
21	Mining	21	2	53,811	7	17	2	28,491	5
31	Food and Tobacco	161	17	141,658	20	214	23	113,085	19
32	Textiles & Apparel	250	26	192,769	27	139	15	119,663	20
33	Wood & Furniture	16	2	7,472	1	24	3	16,324	3
34	Paper & Printing	41	4	16,129	2	38	4	8,115	1
35	Chemicals, Refineries & Rubber	132	14	72,306	10	115	12	59,110	10
36	Pottery & Glass	69	7	33,806	5	75	8	30,656	5
37	Basic Metals	74	8	45,603	6	110	12	50,253	8
38	Metal Fabrication	194	20	134,225	19	184	19	172,797	28
39	Other	3	0	599	0	9	1	2,390	0
40	Electricity	4	0	21,089	3	20	2	7,290	1
	SUM			**719,467**				**608,174**	

We can see that between 1997 and 2014, the share of mining, as well as low technology industries such as food and tobacco, textiles and apparel, paper and printing have been decreasing, and the share of more mid-technology industries such as metals and metal fabrication has been increasing. There are two major exceptions to this rule. The first is that the share of the low-technology industries of wood and furniture have increased in the number of employees between 1997 and 2014. The second is that the share of the mid-technology industry of electricity production has decreased in the number of employees between 1997 and 2014. Apart from these two major changes, however, the number of employees per industry have followed an expected pattern between 1997 and 2014, with the structure of the economy shifting to more mid-size businesses. We can see a more visual representation of this change in Figure 7:

Figure 7: Comparison of the % share of the number of employees in each industry between 1997 and 2014



4.5. New Entries

Like the survival rate of the entire cohort for a given list year, we have also looked at the survival rates of new entries for each year. The results agree with those for the entire cohort: with newer cohorts, the survival rates increase for each duration. There are some notable differences, however, the first being that the data appears to be more periodic with alternating local maxima and minima. This may be due to the fact that the number of new entries seem to be periodic but decreasing as shown on Figure 8. For instance, in 1981 there were 158 new entries to the list from 1980, But this number was 48 in 2012 (the all-time low).

Figure 8: The number of new entries in each yearly cohort for ISO-500

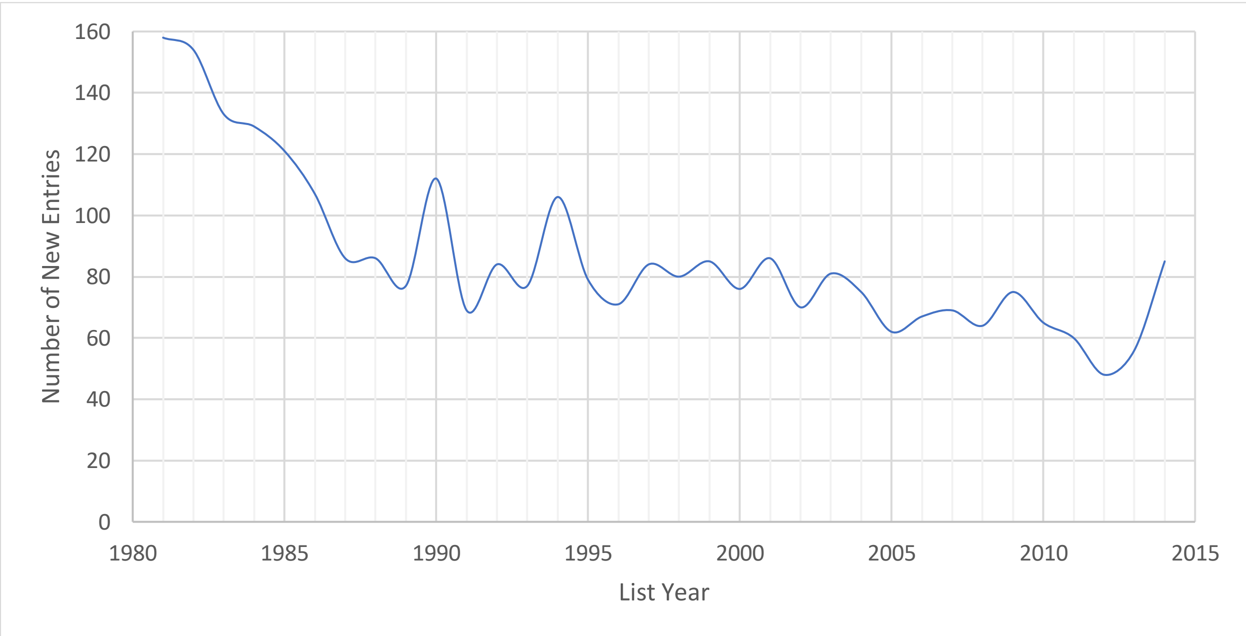
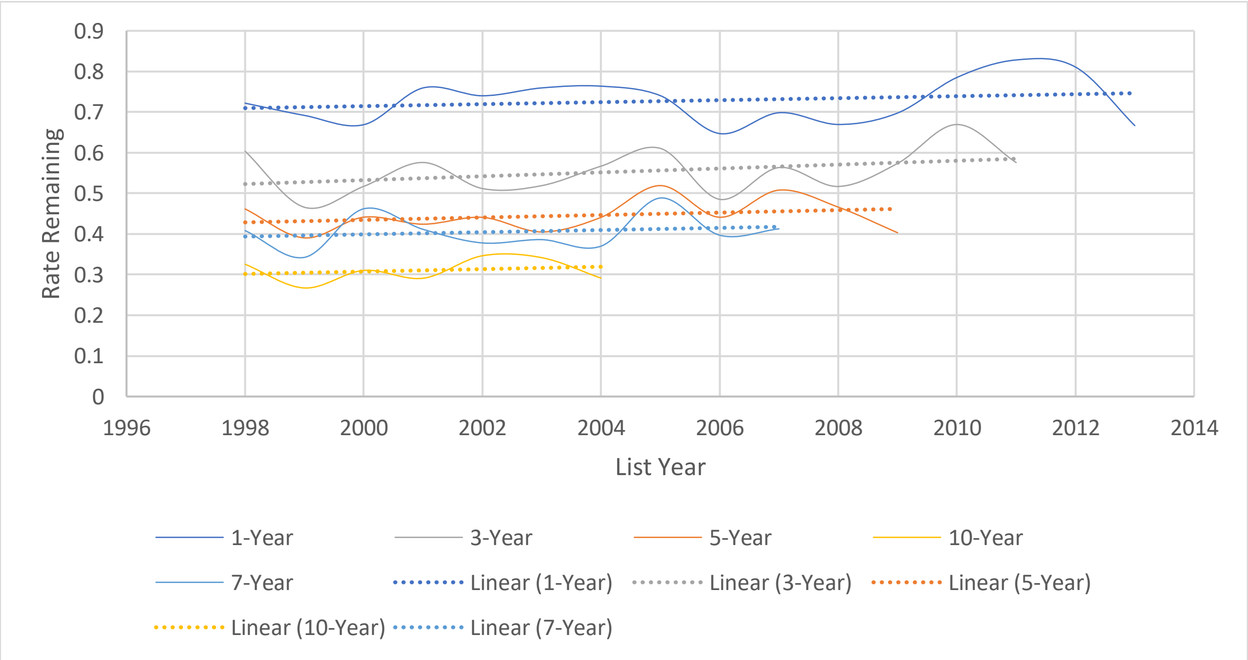


Figure 9 shows the rate remaining in each cohort for the different X-yearly durations. For example, the blue line is the 1-year survival rate across cohorts, which appears to increase. Since the trends are more difficult to visually observe, we have added a trend line to the results.

This pattern is repeated for the ISO-1000 list as shown on Figure 10.

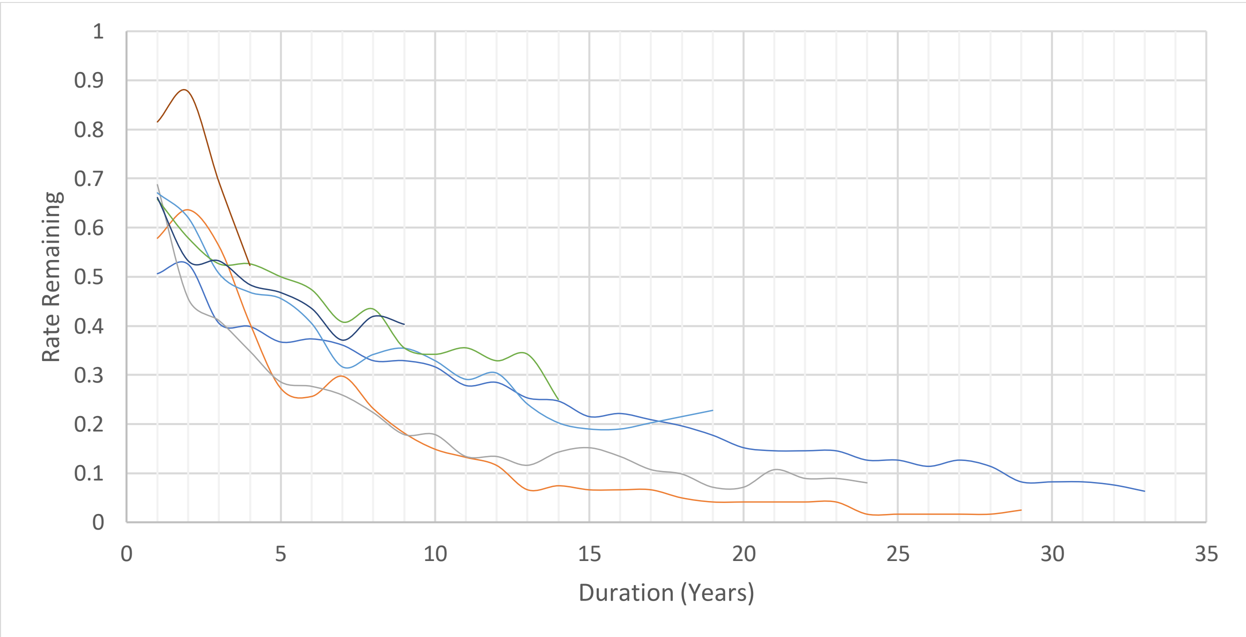
Figure 9: The rate of yearly new entries remaining in the list across every cohort for 5-yearly survival durations in IS0-500

Figure 10: The rate of yearly new entries remaining in the list across every cohort for 5-yearly survival durations in ISO-1000



Unlike the results for the entire cohort, where the survival rates converged as the duration increased for different cohorts, there does not appear to be a pattern in the behavior of survival among new entries for different cohorts (in Figure 10, trend of the survival does not appear to increase or decrease significantly for the different cohorts). This is also shown in Figure 11 for the ISO-500 list. (The ISO-1000 list also does not display any pattern, and therefore has been omitted). Looking at Figure 11, survival appears to drop more haphazardly among the new entries than among the entire list.

Figure 11: The rate of new entries remaining in the list across every year for 5-yearly cohorts in ISO-500



4.6. Crisis Years

We have also looked at the effect of a crisis year in the drop-off rate of firms in ISO-500 and ISO-1000. For this, we have chosen the economic crises in the years 1994 (for ISO-500 only), as well as the 1999 and 2001 economic crises. To gage this, the new entries to the list as well as the full list for any given year has been observed.

For both the full list and among the new entries, (looking at Figure 12 and Figure 13), no particular trend is initially observed. The results for ISO-1000 list (omitted from the paper) are also similar. Of the 106 new entries in 1994, for example, 61% survived for one year. This can be compared to the value in 1993 of 58% and the value in 1995 of 67%.

This is only a partial story, however, since there appears to be a correlation between the number of new entries and crisis years. Looking at Figure 14, there are two points of note. Firstly, in accordance with the increased survival rates of firms in the ISO-1000 and ISO-500 lists, there is a downward trend in the number of new entries over the years. Secondly, if we look at the 1994, 1999, and 2001 crises, the number of new entries in those years seem to be local maxima. For instance, compared to the 77 and 79 entries in 1993 and 1995 respectively, there are 106 new entries in 1994.

Therefore, we can suggest that the crisis years are years of high turnover for the top firms in Turkey, with more new entries coming in than the years immediately preceding and succeeding it, but these years are mostly isolated and do not affect the general downward trend in the number of new entries per year across time. This would make sense, since as there are always 500 (or 1000) members in the ISO lists, a high turnover would imply more firms are leaving the list during the crisis years.

Figure 12: The survival rates of firms in 5-yearly intervals in IS0-500 (with crisis years)

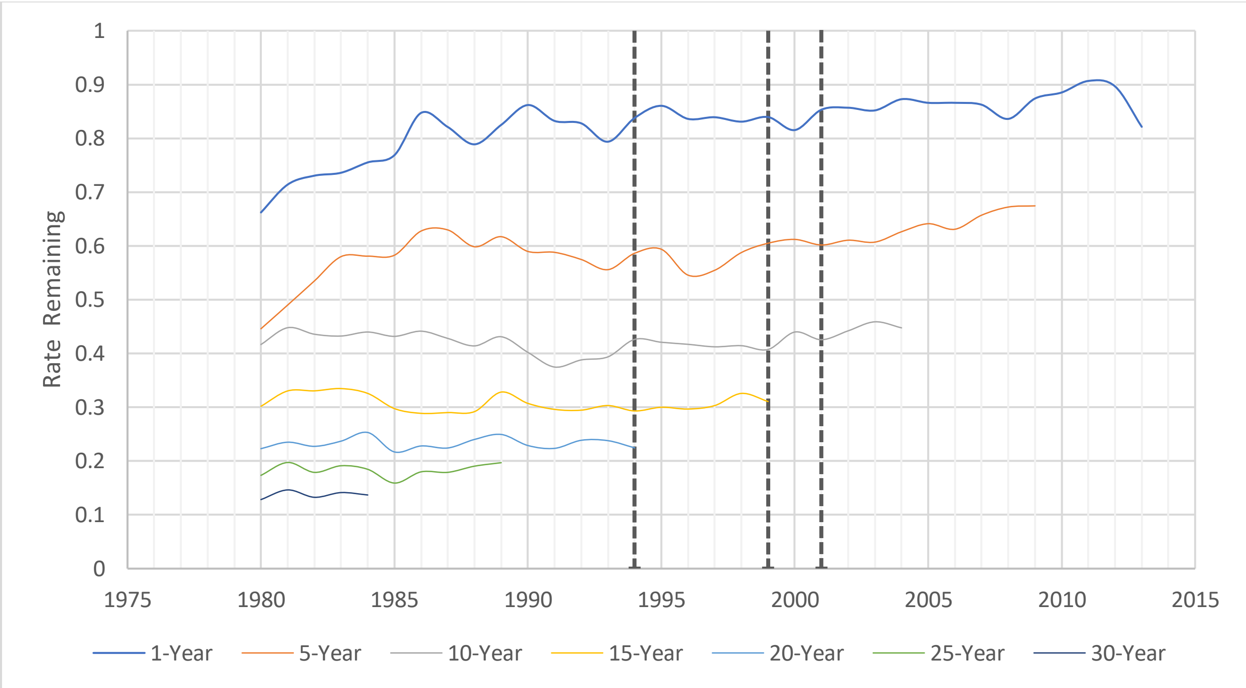


Figure 13: The survival rates of new entries in 5-yearly intervals in IS0-500 (with crisis years)

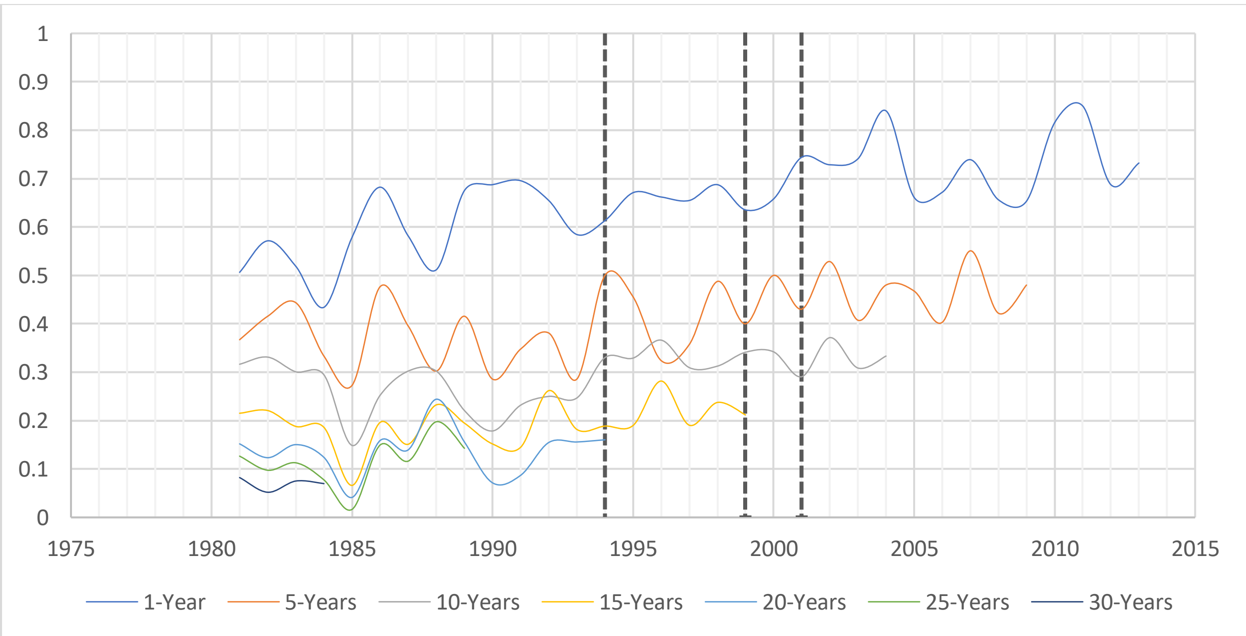
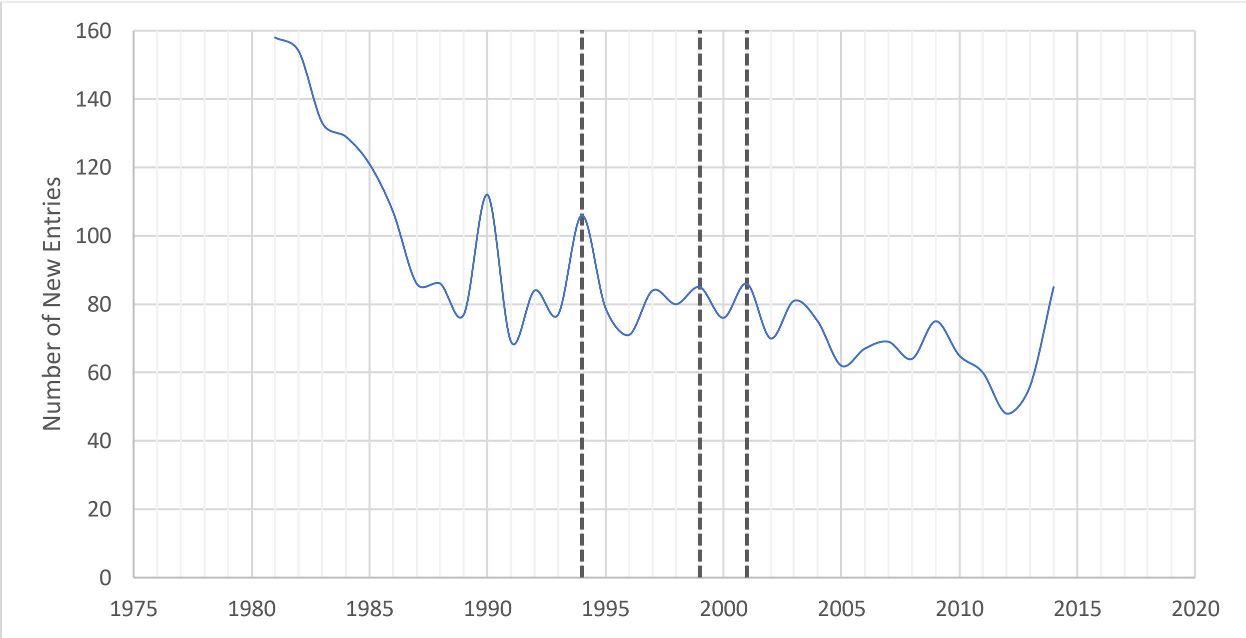


Figure 14: The number of new entries over the years in ISO-500 (with crisis years)



5. MODEL

The data set used in this paper is both right and left censored, using the terminology of Kiefer (1988), meaning that for years before 1980 and after 2014, the observations do not exist for whether the firms are in either list. Several models have been used to estimate the survival of the firms in ISO-500 and ISO-1000 lists, and the chief among them are the hazard and survival models developed in Kaplan-Meier (1958) and Cox (1972).

We denote *T* to be a random variable representing failure time – the time that a firm *i* in ISO-500 or ISO-1000 list will exit the list. Using the terminology of Cox (1972), the survivor function is then defined as:

$$S(t) = pr\{T \geq t\} \tag{1}$$

That is, *S(t)* denotes the probability that failure time for a firm in the list (*T*) has occurred after *t*. For example, S(3) is the probability that a firm will survive (remain in the ISO list) for more than 3 years. From this we produce a hazard function:

$$\lambda(t) = \lim_{\Delta t \to 0+} \frac{pr\{t \leq T < t + \Delta t \mid t \leq T\}}{\Delta t} \tag{2}$$

The hazard function can be interpreted as the age specific failure rate: the rate of event occurrence per unit time as unit time converges to 0. Here, the numerator is the conditional probability that the failure time is in the interval $[t, t + \Delta t)$ given that it did not occur before *t*. The denominator is the length of the interval. By opening the conditional probability:

$$\lambda(t) = \lim_{\Delta t \to 0+} \frac{pr(t \le T < t + \Delta t \mid t \le T)}{\Delta t}$$

$$= \lim_{\Delta t \to 0+} \frac{\frac{pr\{(t \le T) \wedge (t \le T < t + \Delta t)\}}{pr\{t \le T\}}}{\Delta t}$$

$$= \lim_{\Delta t \to 0+} \frac{\frac{pr\{t \le T < t + \Delta t\}}{pr\{t \le T\}}}{\Delta t}$$

$$= \lim_{\Delta t \to 0+} \frac{\frac{pr\{t \le T < t + \Delta t\}}{S(t)}}{\Delta t}$$

For small enough Δt, the numerator can be rewritten as $f(t)dt$, where $f(t)$ is the probability density function of T (i.e. $f(t) = pr\{t = T\}$)and therefore dividing by dt and taking the limit:

$$\lambda(t) = \frac{f(t)}{S(t)} \tag{3}$$

The hazard function at time *t* – the instantaneous rate of the event occurring at time *t* – is the density of events at *t* divided by probability of surviving the event at time *t.* In the case of the firms in the ISO list a hazard rate can be interpreted as the number of firms that exit at time *t* divided by the probability that a firm remains in the list until time *t.*

5.1. Non-Parametric Model

In this paper, the hazard function will be estimated in a few different ways. Firstly, the no initial form for a model is assumed and the hazard is estimated non-parametrically. This is done as per the product-limit estimate first described in Kaplan and Meier (1958). For the product limit estimate, we assume at time t=0, there is a number, *N* of firms in the ISO list. Between t=0 and t=1, some number of firms in the list drop out. This is either due to “failure” or “loss”. A firm fails if it does not turn up on the next year’s list because it no longer is a member of the top 500 or 1000 firms (depending on the list). A firm may be lost if it does not turn up on the next year's list even though it is still a member of the top 500 or 1000 firms. This is due to a firm being unwilling to give its information to that year's survey. For our estimation, these censored observations are not considered as most of the firms that censor their names also do not give some or most of their financial data, which leaves us with little information to work with. We do not believe this would be too much of a problem as for any given year, the number of observations that are thus censored both appear to be randomly distributed and are few compared to the rest of the dataset.

For right-censored data (data cut-off at the last year of the survey), we let *j* be the duration for each spell a firm was in the list (e.g. if a firm was in the list between 1983 and 1994, then *j*=11), h_j be the number of completed spells of duration *j*, and m_j be the number of observations with durations greater than *j*. Then let n_j be the number of spells not completed or not censored before *j:*

$$n_j = \sum^{K} (m_q + h_q)$$

n_j can be interpreted as firms that are at risk at time *j* – the number of firms that have survived before the completion of *j*. For example, among 2000 firms looked at for the length of their survival, $n_1 = 2000$, as all the firms have survived before the completion of *j=1*. A natural estimator for the hazard function, $\lambda(t)$ at discrete time *q=j*, is:

$$\hat{l}(j) = \frac{h_j}{n_j} \tag{4}$$

The survivor function, can then, be estimated as:

$$\hat{S}(j) = \prod_{q=1}^{j} \frac{n_q - h_q}{n_q}$$

$$= \prod_{q=1}^{j} 1 - l(q) \tag{5}$$

This estimator is a special case where the only loss is assumed to be the end censoring. This can be generalized to any losses that occur between periods (accounting for the loss of firms between two years that are not explained by the firm dropping off the list) as per Kaplan and Meier (1958). However, only end censoring is used in this paper.

This can be best illustrated by an example. In Table 3, we have presented a partial table of the ISO-1000 list between the years 2008-2013:

Table 3: New entries into the ISO-1000 list between 2008-2013

	New Entries	**Number of New Entries Remaining after Given Number of Years have Passed**					
		1	**2**	**3**	**4**	**5**	**6**
2008	118	79	59	61	54	55	51
2009	129	90	78	74	69	52	0
2010	121	95	86	81	70	0	0
2011	99	82	72	57	0	0	0
2012	74	60	48	0	0	0	0
2013	84	56	0	0	0	0	0
Sum	625	462	343	273	193	107	51

Here, **New Entries** column denotes the new entries into the ISO-1000 in the given year. Each column after that shows the duration of survival for the new entries. For instance, of the 118 new entries in 2008, 79 have survived after 1 year, 59 have survived after 2. It is important to note that the increase from 59 to 61 between years 2 and 3 show that some firms which had left the list after year 2 have re-entered in year 3. Various methods will be discussed in the coming sections on how to deal with these `survivals`, but for the sake of the example, they will not be considered.

Let us note that there has been a total of 625 new entries into the ISO-1000 list between 2008 and 2013, and the survivors among these new entries are shown in the **Sum** section of the table. Now, let us add Table 4.

Table 4: Hazard and survival function estimate for the partial data between 2008-2013

Years (j)	**h(j)**	**m(j)**	**n(j)**	**l(j)=h(j)/n(j)**	**1-λ(j)**	**S(j)**
1	163	462	625	0.2608	0.7392	0.7392
2	119	343	462	0.257576	0.742424	0.5488
3	70	273	343	0.204082	0.795918	0.4368
4	80	193	273	0.29304	0.70696	0.3088
5	86	107	193	0.445596	0.554404	0.1712
6	56	51	107	0.523364	0.476636	0.0816

- Here, the **Years** column shows the number of years the entries remained in the list.
- **h(j)** column shows the number of completed spell of duration j – i.e. the number of firms that only survived for *j* years. For example, since of the 625 new firms in the list between 2008 and 2013, a total of 462 remain after 1 year, the number of completed spells of duration 1 is 625-462 = 163 (Row 1). Of the remaining 462 firms in year 2, only 343 remain after year 3, meaning, 462-343=119 (Row 2) firms have a completed spell duration of 2 years, etc.
- **m(j)** column shows the number of observations with durations greater than j. This is equal to the sum of surviving firms after year j in Table 3. For example, since, of the 625 new firms in the list between 2008 and 2013, 273 remain after three years m(3)=273, etc.
- **n(j)** column shows the cumulative sum of all the firms whose spell was not censored or completed before j. For example, 625 firms remain in the list remained in the list before the end of year 1, etc.
- We then estimate a hazard function, $l(j)$ shown in the next column.
- Similarly, we estimate a survivor function *S(j)*, shown in the final column.

The non-parametric models will produce a step graph which can be interpreted before using parametric models. This paper will make use of both the Kaplan-Meier estimator for the hazard function as well as an estimated integrated hazard function:

$$\hat{\Lambda}(j) = \sum_{q \leq j} \hat{\lambda}(q)$$

As the sum of all the hazards coming before it, the integrated hazard function retains the memory of the hazards for each specific duration. For example, if the integrated hazard is convex, it would display positive duration dependence of hazard (i.e. hazard of a firm leaving the list increases with duration).

More details on the interpretation of both graphs will be presented with the results.

5.2. The Covariates for the Time Independent Continuous Time Models

Completing the non-parametric analysis of the firms, we try fit different parametrically defined models to the data given a set of firm characteristics. These characteristics are given in the ISO data list.

Before illustrating the model, let us look at the set of covariates which will be used in this paper:

$$\begin{aligned}\mathbf{x}_i = (&LogExitSize, NetworkDummy, EntryProfitability, EntryProductivity, \\ &ExitProfitability, ExitProductivity, ExportDummy, IndustryDummy, CityDummy)\end{aligned} \tag{6}$$

For a vector of covariates $\mathbf{x}_i$.

Throughout this paper, we shall use the Proportional Hazards assumption, which very simply means that the covariates $\mathbf{x}_i$ increase the hazards proportionally to the baseline. In other words, if $\lambda_0(t)$ represents the baseline hazard of a firm dependent on *t* but not on any $\{x \in \mathbf{x}_i\}$, then the firm specific hazard function at time *t* is assumed to be follow:

$$\lambda(t, \mathbf{x}_i) = \lambda_0(t)\psi(\mathbf{x}_i) \tag{7}$$

where, $\psi(\mathbf{x}_i) = \exp(\mathbf{x}_i'\beta)$ is the firm-specific function of covariates $\mathbf{x}_i$, the distributions of which will be further specified below. This allows us to separate the time component of a firm's hazard out of the firm dependent covariates $\mathbf{x}_i$ – i.e. the covariates themselves are not time-dependent.

This assumption allows us to propose continuous time models but also limits us, in that we have to use time-invariant covariates for each firm. Therefore, using this assumption, only certain covariates such as the size of a firm on the year of its entry or its exit can be modeled, and not the full extent of the data which we have the size of a firm on each year of its inclusion in the list. This can be seen in the set of covariates in **(6)**, which only models entry and exit year covariates for size (number of employees), productivity, profitability. Similarly, the exporter status has been given a firm if it has exported for at least one year during its presence in the list, if more than one industry has been specified over the years, the one that has been specified for the most years has been used, and a similar method has been employed for the city dummy as well. This method will be altered using another specification, a time variant, discrete-time model. In the results section of this paper, we present some tests to see if the proportional hazard assumption holds.

5.3. Parametric Model

The non-parametric model will form the basis of a parametric approach to the hazard function. This has been first described in Cox (1972) and later in Kiefer (1988). In the parametric model, a certain shape for the hazard function is assumed and then the functional parameters are estimated. Functional forms with the following distributions have been estimated: exponential, Weibull, Gompertz, Log-logistic, as well as a semi-parametric model (Cox Hazard Function). While the detailed specifications of each of these models will be further discussed, in general, the purpose of each of these models is to compare the hazard ratios between different groups among the firms. Given a baseline hazard following a certain functional form, what effect does a set of covariates (**6**) have on the hazard functions? Before further analyzing each of the different hazard functions following the different distributions, let us first show the interpretation of the hazard ratios for the generalized model and state the Survival function. Rephrasing (**Eq. 7)**:

$$\lambda(t, \mathbf{x}_i) = \lambda_0(t)\exp(\mathbf{x}_i'\beta)$$

Then, at time $\bar{t}$ for two firms, ***j*** and ***k*** with the covariates $\mathbf{x}_j$ and $\mathbf{x}_k$, we have:

$$\frac{\lambda(\bar{t},\mathbf{x}_j)}{\lambda(\bar{t},\mathbf{x}_k)} = \frac{\lambda_0(\bar{t})\exp(\mathbf{x}_j'\beta)}{\lambda_0(\bar{t})\exp(\mathbf{x}_k'\beta)}$$

$$= \exp(\mathbf{x}_j'\beta - \mathbf{x}_j'\beta)$$

$$= \exp(\beta[\mathbf{x}_j' - \mathbf{x}_k'])$$

Then,

$$\log\left(\frac{\lambda(\bar{t},\mathbf{x}_j)}{\lambda(\bar{t},\mathbf{x}_k)}\right) = \beta[\mathbf{x}_j' - \mathbf{x}_k'] \tag{8}$$

As the right-hand side of the of the equation does not depend on time (by assumption), the proportional difference in hazards is constant.

Now, we can modify **equation 8** for two firms, holding equal all their characteristics but the variable *v* for both *j* and *k:* x_{jv} and x_{kv}:

$$\frac{\lambda(\bar{t},x_{jv})}{\lambda(\bar{t},x_{kv})} = \exp(\beta_v[x_{jv} - x_{kv}]) \tag{9}$$

In proportional hazards models, the coefficient β_v on the *v*th covariate has the property:

$$\beta_v = \frac{\mathrm{d}\log\lambda(t,\mathbf{x})}{\mathrm{d}\,x_k}$$

Therefore, the hazard ratios can then be interpreted as the proportional effect on the hazard of the absolute change in the corresponding covariate.

For instance, in the results section, one of the coefficients for the log of exit size of a firm (the number of employees the year of its exit from the ISO list), is -0.309.

The coefficient can be translated to the hazard ratio by: $\exp(-0.309) = 0.734$. Multiplying the number of by e reduces the hazard by 27%. This is the elasticity of the covariate but is not particularly meaningful to interpret.

Therefore, to find the effect of 1% increase in the number of employees in the exit year on the relative hazard, we can use **equation 9** to give:

$$\lambda(t, \%\#\text{employees} = 1\%) = \exp(-0.309 \log(1.01)) = 0.9969$$

Hence, a 1% increase in the number of employees reduces the hazard of exit by 0.3%.

For a dummy variable, the interpretation is much simpler. Modifying **equation 9** for dummy variables $x_{jv} = 1$ and $x_{kv} = 0$, we get:

$$\frac{\lambda(\bar{t}, 1)}{\lambda(\bar{t}, 0)} = \exp(\beta_v)$$

Here, the hazard ratio can be simply interpreted as the change in hazard between the control and test groups. For example, in the results section, for the export dummy which takes on 1 if the firm is an exporter, 0 otherwise, one of the regressions gave a coefficient of -0.45, which translates to a hazard ratio of 0.63. This mean that a firm which is an exporter has 63% of the hazard (to leave) of a non-exporter.

The Survival function of this model can be stated as:

$$S(t, \mathbf{x}_i) = \exp\left[-\exp(\mathbf{x}_i'\beta) \int_0^t \lambda_0(t)dt\right] \tag{10}$$

The difference in each of the models are the assumptions made about the shape of the baseline hazard function. The most strict of these assumptions is with the exponential hazard function, and they are relaxed with the Weibull and Gompertz functions.

We will see the outcomes on the *Results* section of this paper. However, while we will present some of the outcomes across models, we will only give a detailed analysis using the Weibull Model. This model fits our assumptions and our non-parametric analysis the best. The specific limitations and assumptions of each model will be presented below but here, we can state that the Weibull model’s assumption of a monotonically changing hazard (in our case, decreasing) will be correct. This makes the most intuitive sense, as there is a good chance that longer a firm stays in the list, the less it is likely to leave the list, as the major firms in the lists will more likely become `fixtures` and therefore will be more difficult to displace.

5.3.1. Exponential Distribution:

The first functional form to be assumed for the ISO data will have the exponential distribution, where the survival and hazard functions are assumed to have the forms:

$$S(t) = \exp(-\gamma t) \tag{11a}$$

$$\lambda(t) = \gamma \tag{11b}$$

The exponential function assumes that the hazard in each period (γ) is constant, and therefore, the length of the previous survival of a firm (i.e. how long it has stayed in the list) has no effect on its risk of dropping out. Whether or not this assumption is true will depend on the specific shape that the non-parametric estimation will take: whether the graphs of the estimated survivor and the integrated hazard functions will be straight, convex or concave.

5.3.2. Weibull Distribution:

We will also assume that the ISO data will have the Weibull Distribution:

$$S(t) = \exp(-\gamma t^{\alpha}) \tag{12a}$$

$$\lambda(t) = \gamma \alpha t^{\alpha - 1} \tag{12b}$$

The Weibull distribution is a generalization of the exponential distribution which is a special case, where $\alpha = 1$. The α coefficient (which is also to be estimated) rescales the time axis to describe the behavior of the hazard function, where the hazard is thought to increase with time if $\alpha > 1$ and decrease with time if $\alpha < 1$.

5.3.3. Log-Logistic Distribution:

Related to the Weibull distribution is the log-logistic distribution, with the hazard function:

$$\lambda(t) = \frac{\gamma \alpha t^{\alpha-1}}{1 + t^{\alpha}\gamma} \tag{13}$$

The numerator is the same as the hazard function with Weibull distribution. The denominator, on the other hand, adds non-monotonicity to the hazard. For coefficients $\alpha \in (0,1]$, the hazard function decreases with duration (for $\alpha > 1$, it first increases and then decreases).

5.3.4. Gompertz Distribution:

The Gompertz distribution assumes that the failures are described by geometric progression. The failure time *T* will be assumed to follow a Gompertz distribution with two strictly positive parameters *a* and *b* if:

$$S(t) = \exp\left(-\frac{a}{b}\left(e^{bt} - 1\right)\right) \tag{14a}$$

$$\lambda(t) = a\exp(bt) \tag{14b}$$

This distribution was initially developed in order to study life tables and therefore most accurately fits data which increase in risk geometrically with time. One can think of the risk of death for an individual between the ages of 20 and 30 compared to the risk of death between 70 and 80: the risk increases geometrically over time with the latter group having much higher risk of death compared to the former group. This particular distribution would fit a model whose non-parametric estimation of hazard appears to increase with time and would allow us to model the tail of the graph with increased accuracy (a convex graph for the integrated hazard).

5.4. Time Dependent Covariates

As is discussed in the *Data* section of this paper, the datasets that we work with in this paper – namely ISO-500 and ISO-1000 lists, have values for firm size, value from sales, export amount, total capital and profit/loss for each individual year from which we can calculate values for profitability and productivity on a yearly basis. This would mean that for each firm we will have multiple values for profitability and productivity as well as firm size that vary discretely with time. Rodriguez (2007) notes that "time-varying covariates … [have] rarely been done in applications" (p.14) For the case of our study, the covariates that change discretely with time include the number of employees in the firm, the values for the profits and revenue. As, we can no longer hold the assumption that the covariates $\mathbf{x}_i$ of a firm i are time-invariant, we can no longer separate out the firm specific covariates from the time-dependent baseline hazard.

Primarily, **equations (a.2) and (a.3)** (from the Appendix) is modified into:

$$\lambda(t, \mathbf{x}_i(t)) = \lambda_0(t)\exp(\mathbf{x}_i(t)'\beta) \tag{15}$$

$$S(t, \mathbf{x}_i(t)) = \exp\left[-\int_0^t \lambda_0(t)\exp(\mathbf{x}_i'(t)\beta)dt\right] \tag{16}$$

where some or all of the of the $\mathbf{x}_i$ now may take different values across time. As we are specifying values for $\mathbf{x}_i$ across time, we can no longer factor the survivor function, but we can assume that each covariate is constant within a given time interval and calculate it piecewise. We will follow the derivation by Jenkins and first let q_j be the date making the end of the time interval $(q_{j-1}, q_j]$.

With the assumption of a constant $\mathbf{x}_i$ within the period q_j, we can rewrite the survival function (12) as:

$$\begin{aligned} S(q_j, \mathbf{x}_i) &= \exp\left[-\int_0^{q_j} \lambda_0(q)\exp(\mathbf{x}_i'\beta)dt\right] \\ &= \exp\left[-\exp(\mathbf{x}_i'\beta)\int_0^{q_j} \lambda_0(t)dt\right] \end{aligned}$$

$$= \exp\left[-L_j \exp(\mathbf{x}_i'\beta)\right]$$

where, $L_j \equiv \int_0^{q_j} \lambda_0(t, \mathbf{x}_i)$. Next, we can define a discrete time hazard function h_j at t_j as:

$$l(\mathbf{x}_i) = \frac{S\left(q_{j-1}, \mathbf{x}_i\right) - S(q_j, \mathbf{x}_i)}{S(q_{j-1}, \mathbf{x}_i)}$$

$$= 1 - \exp[\exp(\mathbf{x}_i'\beta)(L_{j-1} - L_j)]$$

Rearranging gives us:

$$\log\left(-\log\left[1 - l_j(\mathbf{x}_i)\right]\right) = \mathbf{x}_i'\beta + \log\left(L_{j-1} - L_j\right) \quad (17)$$

We can also get a discrete time baseline hazard function h_{0j} similarly, giving us:

$$\log\left(-\log\left[1 - l_{0j}\right]\right) = \log\left(L_{j-1} - L_j\right)$$

$$= c(j) \quad (18)$$

Combining **(17) and (18)** and rearranging, we can get:

$$l_j(\mathbf{x}_i) \equiv l\left(q_j, \mathbf{x}_i\right) = 1 - \exp(-\exp(\mathbf{x}_i'\beta + c(j)) \quad (19)$$

Which is the complementary log-log model. Recalling **(4)** which estimates the survival rate for discrete time q, at any period j, the hazard rate will be.

$$S_j = \prod_{q=1}^{j} 1 - l(q)$$

$$= [1 - l(1)][1 - l(2)] \ldots [1 - l(q)]$$

Using (9) and (10) we will be able to predict the hazard rates across time by assuming a form (or numerous forms) for the baseline hazard function $c(t)$ and fitting a complementary log-log function to the data. Then we will predict the hazard ratios for different firms at any given period.

In our model, we first define a `run` variable. This variable takes the value 1 for the first year a given firm appears on the list, and then increases by 1 until the firm drops out of the list, or the censoring event happens (i.e. the list years hits 2014).

Next, we define the binary 'event' variable. If the firm has left the list before 2014, then the 'event' variable takes the value '1' for its last year in the list. Otherwise, it takes the value '0'. If the firm's last year on the list is 2014, then 'event' takes the value 0.

There is the issue of a firm that has entered and left the list and re-entered during the 34 years under study. The best way to explain what values 'event' and 'run' variables take is by an example.

For the firm named, "ALTINBAS MUCEVHERAT IMALATI VE DIS TICARET A.S." in ISO-500, first let us set up a value table, followed by an explanation:

Name	**Year**	**Run**	**Censored**	**Event**
ALTINBAS…	1995	1	0	0
ALTINBAS…	1996	2	0	0
ALTINBAS…	1997	3	0	1
ALTINBAS…	2012	1	1	0
ALTINBAS…	2013	2	1	0
ALTINBAS…	2014	3	1	0

This particular set-up for the time-dependent covariates has been described by Prentice and Gloecker (1978) but has been expanded in more practical detail by Jenkins (2007).

Here, the firm had two three year runs: 1995-1997 and 2012-2014. The first run ended when the firm left the list, and hence the 'event' variable takes on the value 1 in its final year. In its 2012-2014 run, the list ends before we know whether the firm survived or not, hence the 'censored' variable takes on the value 1 and 'event' takes on the value 0 for all years that it has been in the list.

Then we set up a baseline hazard function for the firm with various shapes, referred to as h_0. These have been: $run, run^2, run^3, \ln(run)$. These are different specifications on how the baseline hazard of a firm changes with time, and are chosen among a large selection of possible shapes. run assumes that the baseline hazard changes linearly with time, run^2 and run^3 assume that it changes as either a quadratic or cubic polynomial, and finally $\ln(run)$ assumes it changes with log time.

Finally we run a complementary log-log regression of the 'event' variable on the following covariates:

$$\begin{aligned} x_i = \{ & secular, religious, \ln(size), export, Productivity, Profitability, \\ & IndustryDummy, CityDummy\} \end{aligned} \tag{20}$$

Most of these covariates have been defined similarly to **those in (6).**

One issue that can be taken up with the above approach is that it treats the same firm as two different firms for the two different durations. In fact, the model does not take into account the firm at all, instead, defining it through the covariates for its industry, the city where it is located, and the productivity and profitability changing with its run. This is different than the time invariant model which had one set of values per firm. However, that model took for the entry and exit years of a firm (and the *spell* it was in the list) two different specifications. For a given firm, it either took its longest run as the *spell,* or its total run. For ALTINBAS above, this gave a *spell* value of '3' or '6' depending on the way spell was specified. Both definitions had their problems. If it took the longest run, then it discarded some of the data, and if it took the total run, the entry and exit year size, productivity and profitability values would be skewed. If ALTINBAS was not censored (as its final year was 2014), this would have meant that a 6-year spell was defined with the entry profit and revenue values for 1995 and exit ones for 2014.

In practice, these do not make much difference. Both specifications give very similar results in the survival analysis both in values and the signs of the coefficients. However, both methods have their drawbacks, and this will allow us to compare with the results we had for the time-invariant model.

6. RESULTS AND DISCUSSION

The analysis used on the data is survival analysis. Here, “the event” has been described as the firm leaving the given list, and the time until the event has been defined in two ways: (1) the longest spell length, and (2) the total spell length. For most of the firms in the list the two specifications are identical. There are however, some firms which have entered the list in a given year, and then left the list, only to enter it again at a later date. For these firms, (1) defines the longest time a given firm has spent in the list and (2) defines the total time spent in the list. For example, for the example given in the model section, ALTINBAS, the longest run of the firm is 3 years, and the total run of the firm is 6 years. Since the firms that leave the list and return to it at a later date are a minority, and the two specifications differ very little both qualitatively and quantitatively, only the results for specification (2) are shown for the rest of this paper.

6.1. Non-Parametric Analysis

The non-parametric estimation for the survival of firms in the ISO-500 and ISO-1000 lists are shown in Figure 15 and Figure 16. These are the graphical estimates of the survival function $\hat{S}(q)$ described in **equation (5)**.

Figure 15 shows the yearly survival rate estimates for all the years between 1980-2014 for the ISO-500 list, and we can see that the yearly survival of a cohort drops to just below 50% at the 5^{th} year, with the shape of the function being concave in the following years.

Figure 16 shows the yearly survival estimates between 1997-2014 for the ISO-1000 list, and again, the yearly survival for a cohort drops to just below 50% on the 5^{th} year. Also, we can see that for both graphs, the 15-year survival is between 20-30%, with the survival in ISO-500 list being slightly lower.

Figure 15: Kaplan-Meier survival estimate for the ISO-500 list

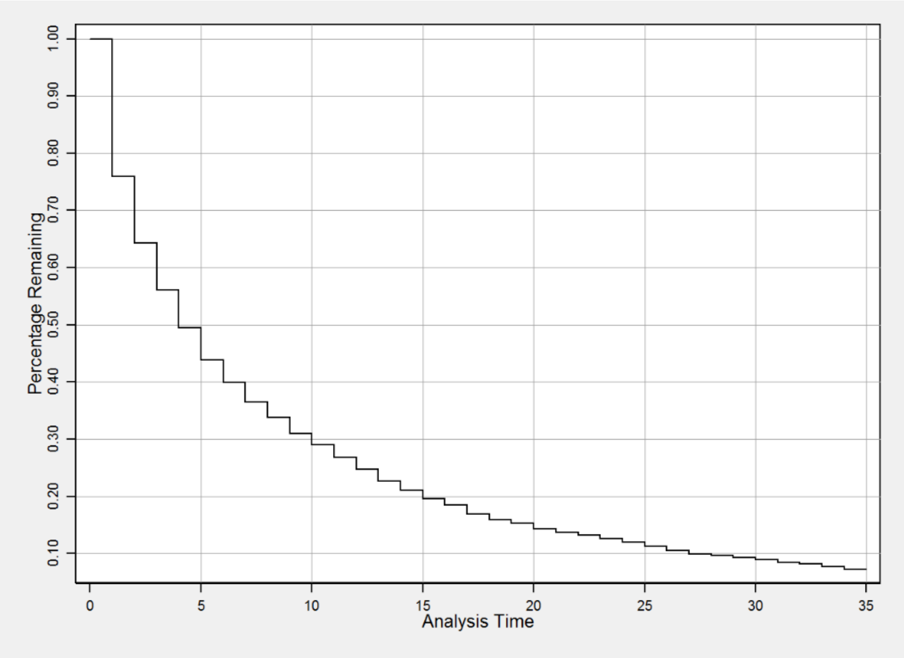


Figure 16: Kaplan-Meier survival estimate for the ISO-1000 list

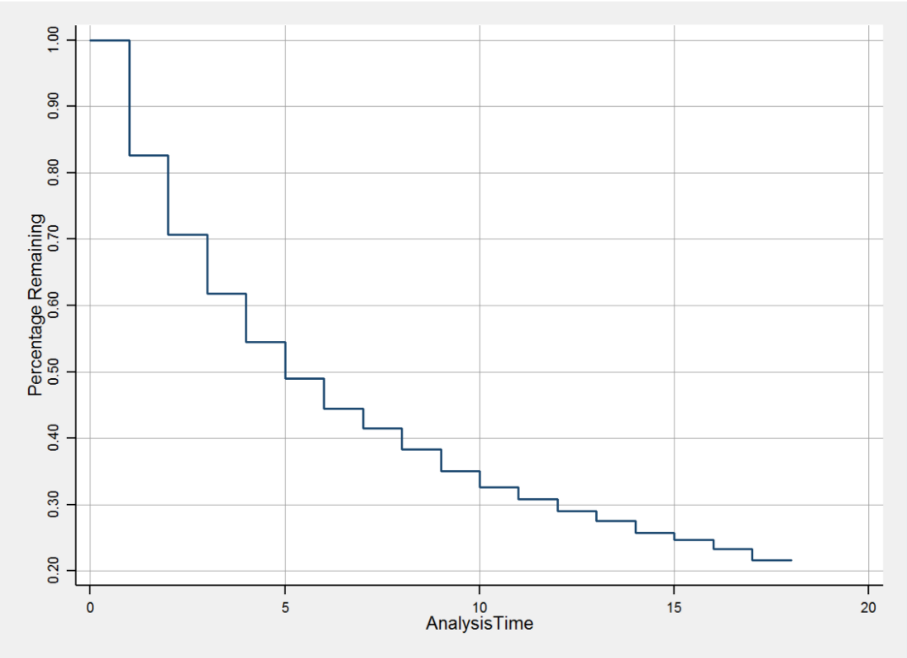
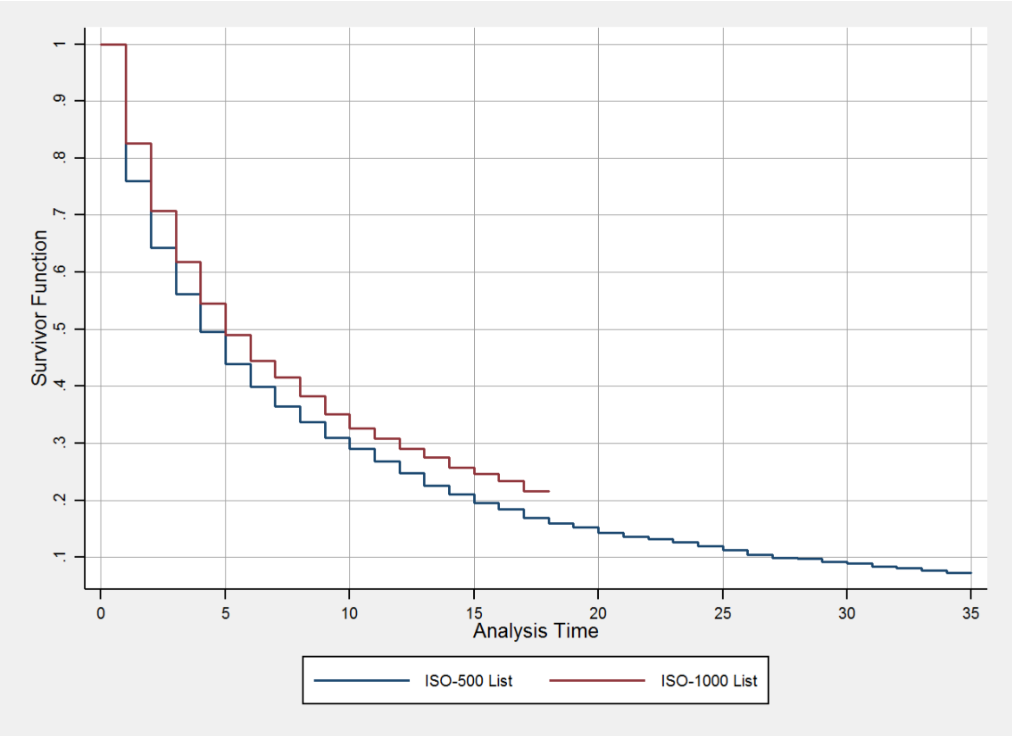


Figure 17 shows that the survivor function for the ISO-1000 list is slightly higher than that for the ISO-500 list. The top 1000 firms are more stable and less likely to change, therefore, than the top 500. As a larger list, the ISO-1000 seems to be less volatile than the smaller one.

Figure 17: Kaplan-Meier survival estimate for the ISO-500 list vs ISO-1000 list



In Figure 18 and Figure 19, we have the cumulative hazard estimate, $\Lambda(q)$. As the cumulative sum of the yearly hazards for each list, it shows us the behavior of the hazard function over time.

The cumulative hazard estimate for the ISO-500 list is slightly concave meaning that as the years a firm remains in the list increases, the hazard drops. This makes intuitive sense as well, since a firm that has consistently remained as one of the top 500 firms in Turkey would be more likely to remain so than a firm that only entered that year. However, the graph is only slightly concave, and therefore, a parametric estimation with an exponential distribution (which assumes a straight cumulative hazard curve) might also fit the data.

The cumulative hazard estimate for ISO-1000 is more linear than that for ISO-500: time has a smaller effect on the hazards in the larger list. Given that Table 19 shows us that the hazards for the larger list are lower overall, we can deduce that the firm survival in the larger list is both higher and more consistent than the in the smaller one.

Figure 18: Cumulative hazard estimate for the ISO-500 list

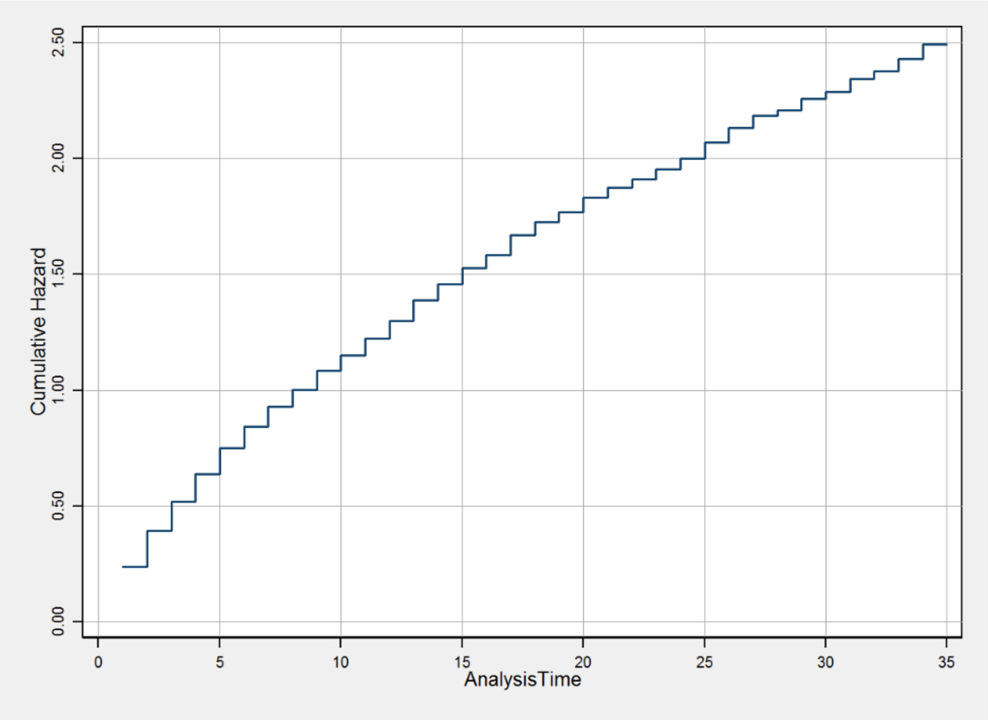
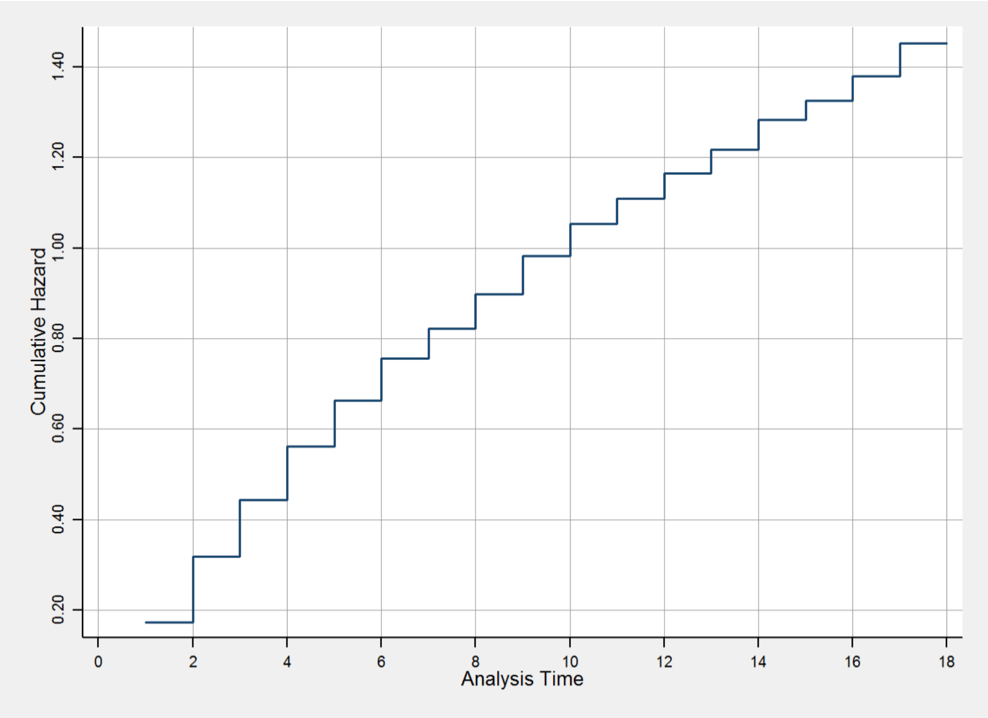


Figure 19: Cumulative hazard estimate for the ISO-1000 list



One possible explanation of this is that the Sales values of the firms, which the lists are based on, drop exponentially.

Figure 20 shows the Average Sales Values of the top 1000 firms in Turkey between 1997-2014. An average sales value has been taken over the time period for the firms occupying each ranking position and plotted. It shows that there is a very large drop in the sales figures of the firms in the first few hundred positions, but this starts to follow a more linear trend in the bottom 500 (as shown in Figure 21). Therefore, for a firm occupying the top positions (down to around the 200th position), to drop-off from the list, they would need to make a very small fraction their sales. For example, a firm at the top of the list, **TUPRAS** which had sale of around ₺40,000,000,000 would need to make sale of less than around ₺213,000,000 (around 0.5% of its 2013 sales) in 2014 in order to drop from the ISO-500 list. This can be compared with **Gamateks Tekstil San. ve Tic. A.S.**, the firm at the 501st place in the 2013 list, which made around ₺188,000,000 in 2013, and would need to make less than around ₺95,000,000 (around 51% of its 2013 sales) in order to drop from the ISO-1000 list. Therefore, the top firms would need to lose a lot more of their sales than their smaller counterparts, making the less linear top half more resilient.

Figure 20: The average real sales values for the ISO-1000 firms between 1997-2014 versus their list ranking

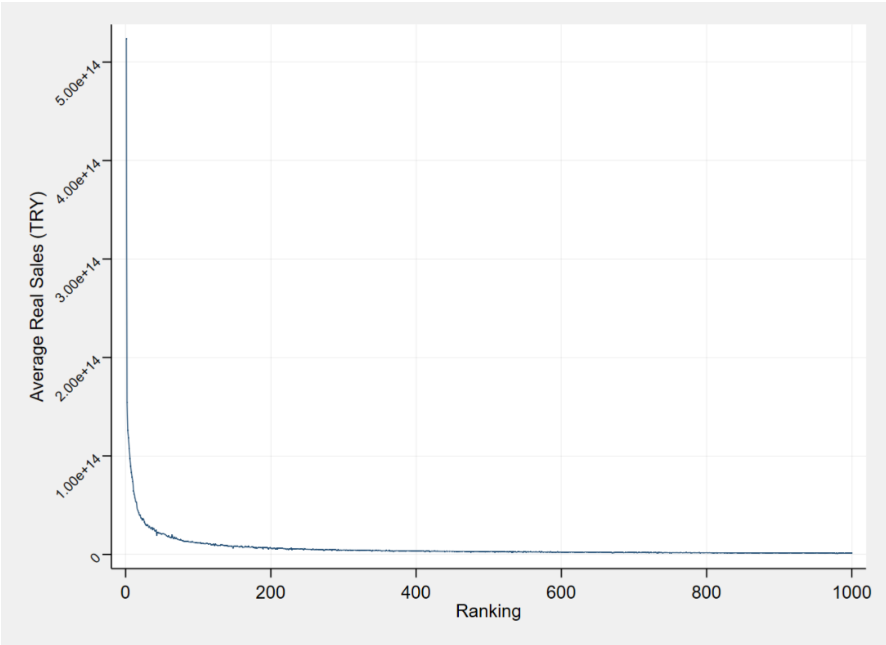


Figure 21: The average real sales values for the bottom 500 firms in the ISO-1000 list between 1997-2014 versus Their list ranking

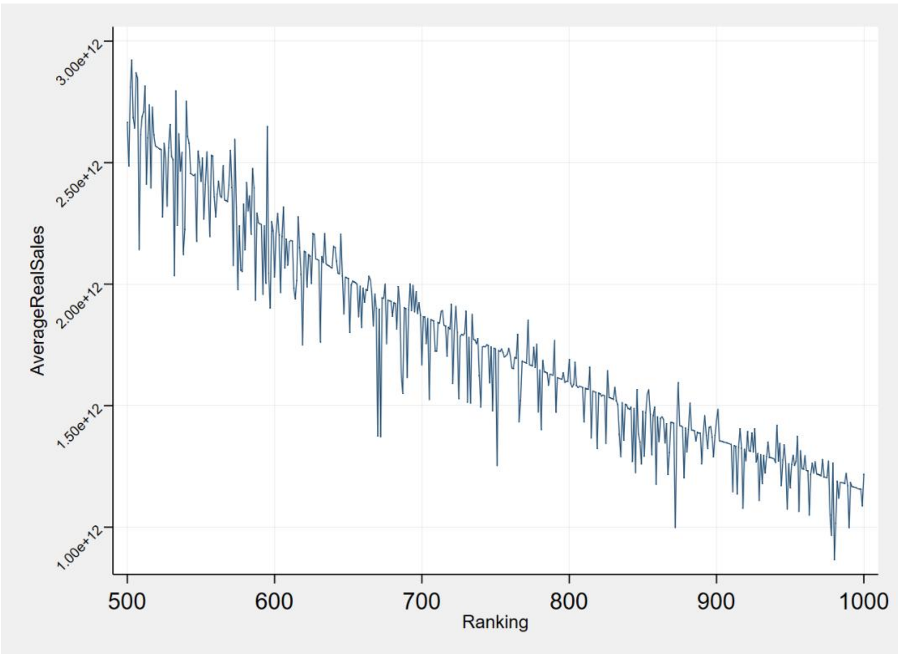
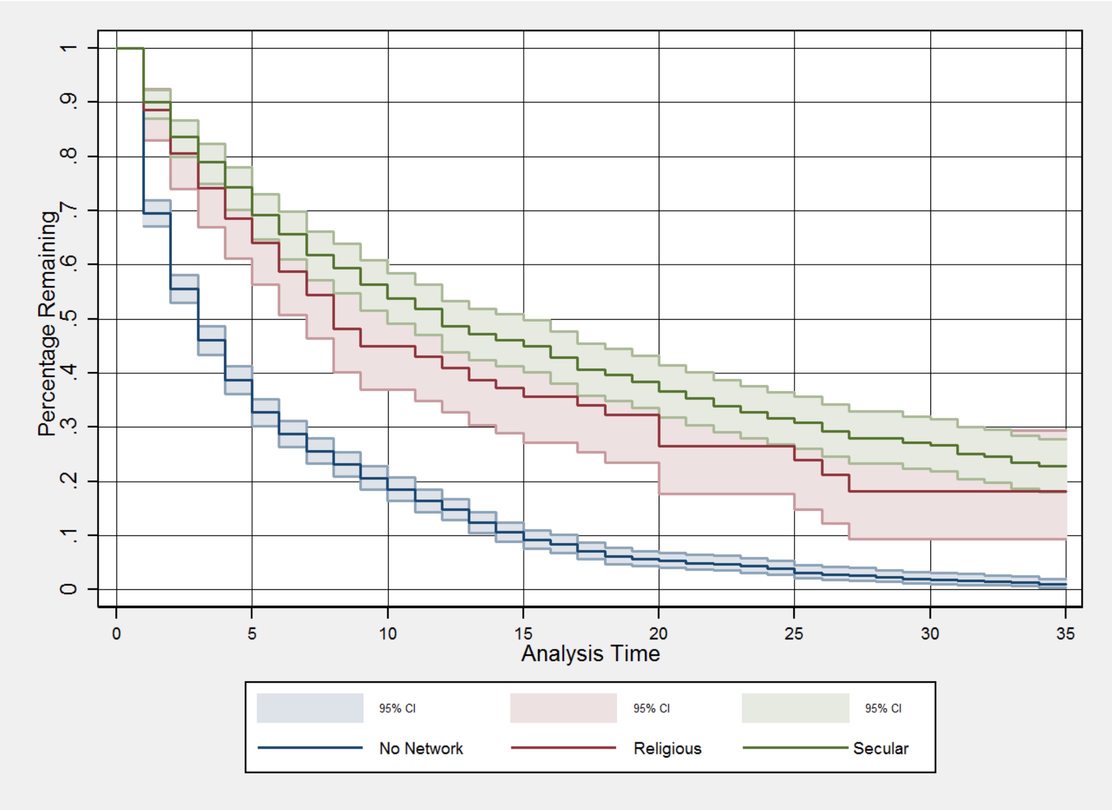


Figure 22 shows the Kaplan-Meier survival estimate for the ISO-500 firms by their affiliation in a religious or secular business network or no network with the 95% confidence intervals displayed around the graphs. The way that these networks were constructed, and assumptions are in the data section of this paper.

Firms in secular networks consistently perform better than firms in either religious network or unaffiliated firms. Within the 95% confidence interval, however, there is a limited overlap between the secular and religious networks, whereas both perform unambiguously better than the unaffiliated firms. Even with this overlap, secular networks perform better than religious networks. This agrees with our initial observations in the descriptive statistics section of the paper in Figure 6, where we saw the average drop off rate of the firms in the ISO-1000 list being higher in the unaffiliated networks, followed by the religious networks, and finally the secular networks. The graph for ISO-1000 is similar and therefore has been omitted.

This is again in agreement with the initial statistics, where the yearly performance of the affiliated firms was consistently higher than the non-affiliated ones, and those in secular networks performed better than the religious networks.

Figure 22: Kaplan-Meier survival estimate for the ISO-500 list by network affiliation



6.2. Parametric Analysis

Before we delve further into the hazard rates for the various different models, we will present in Table 5 the percent changes associated with the changes in the continuous covariates. As discussed in the Model section, the hazard rates for a log-log regression is not necessarily easy to interpret. In Table 5 we look at the change in survival, given 1% increase in the presented covariates with all the controls. For example, a 1% increase in the number of employees in a firm's last year in the list, decreases the probability of the event (i.e. exit) by 0.4%. Also as stated in the Model section of this paper, an in-depth discussion for the hazard ratios will be presented for the Weibull model only but we will present the effect of the dummies as well as a for the change in the various covariates across models in Table 5-Table 7 and give a summary of all models in Table 16 at the end of this section.

Table 5: Percent changes in the survival with 1% increase in the continuous covariates (green statistically significant)

	ISO-500				ISO-1000			
1% Increase in:	**Exponential**	**Weibull**	**Log-Logistic**	**Gompertz**	**Exponential**	**Weibull**	**Log-Logistic**	**Gompertz**
# Emp. Before the Exit Year	-0.40	-0.55	-0.22	-0.52	-0.49	-0.68	-0.40	-0.62
Entry Productivity	0.02	0.06	-0.07	0.05	-0.22	-0.35	-0.18	-0.29
Entry Profitability	0.17	0.23	0.24	0.24	0.26	0.39	0.25	0.34
Exit Productivity	-0.14	-0.19	-0.11	-0.19	-0.17	-0.21	-0.10	-0.20
Exit Profitability	0.01	0.03	-0.07	0.02	-0.04	-0.05	-0.07	-0.05

The results across models seem to agree: the values for the entry profitability, exit profitability (The profitability of the firm on its final year in the list), as well as the number of employees all are positively correlated with survival. There are some issues with the entry productivity which will be discussed further in this section.

Next, before the discussion, we will present the statistically significant dummies for each model in Table 6 and Table 7. None of the dummies presented are negatively correlated with survival. In both lists, among the industry dummies, Industrial Chemicals and Metals industries is positively correlated with survival across all models, followed by Coal Mining, which appears to be positively correlated with survival across all models, except for the exponential one in both lists. There are some differences between the ISO-500 and ISO-1000 lists, however. First, in the larger list, apparel, beverage, iron and steel, scientific equipment and textiles industries are positively correlated with firm survival, whereas in the smaller list, they are not. This implies that of the top 1000 firms in Turkey, among the largest ones (with the most sales), firms that produce consumer goods do not survive as well as the heavy industries.

Table 6: The statistically significant industry dummies by model

	ISO-500				ISO-1000			
	Exponential	*Weibull*	*Log-Logistic*	*Gompertz*	*Exponential*	*Weibull*	*Log-Logistic*	*Gompertz*
Industry								
Apparel (No Footwear)								
Bevarage Industries								
Chemicals								
Coal Mining								
Electrial Machinery								
Furniture (No Metal)								
Footwear								
Glass								
Industrial Chemicals								
Iron and Steel								
Leather Products (No Footwear)								
Metal (No Iron)								
Metal Products (No Machniery)								
Mineral Products (No Metal)								
Other Manufacturing								
Paper								
Petroleum Coal (Misc)								
Plastic								
Printing Publishing								
Rubber								
Scientific Equipment								
Textiles								
Transport Equipment								

Among the city dummies, Ankara, Balikesir, Gaziantep, Istanbul, Izmir are highly correlated with firm survival across all firms in both lists, followed by Kayseri in statistical significance. This is not surprising as these are major cities and industrial centers in Turkey. There are some interesting observations to make in comparing the two lists. First, firms in Adana, Kutahya and Tekirdag are more likely to survive in the larger list, but there is no correlation in the smaller one. For this, we can look at the breakdown of industries in these cities. For Adana, almost 41% of the manufacturing firms in the city are textile firms. As we have seen with the industry dummies, smaller textile firms seem to survive longer than larger ones. Therefore, most of Adana's firms would perform better in the ISO-1000 list. Similar with Tekirdag, the largest share of manufacturing firms operating out of the city are Textile firms with 26% of the firms belonging to that group.

With Kutahya, almost 74% of the manufacturing firm there are Coal Mining firms, but among the top 1000, 14% of them are textile firms, while there are no textile firms in Kutahya among the top 500.

Table 7: The statistically significant city dummies by model

	ISO-500				**ISO-1000**			
	Exponential	*Weibull*	*Log-Logistic*	*Gompertz*	*Exponential*	*Weibull*	*Log-Logistic*	*Gompertz*
City								
Adana								
Adiyaman								
Ankara								
Antalya								
Balikesir								
Bursa								
Denizli								
Duzce								
Elazig								
Eskisehir								
Gaziantep								
Giresun								
Hatay								
Istanbul								
Izmir								
Kahramanmaras								
Karabuk								
Kars								
Kastamonu								
Kayseri								
Kutahya								
Kocaeli								
Konya								
Kutahya								
Manisa								
Samsun								
Sivas								
Tekirdag								

As a reminder, the rephrased equation **(7)** is the general form of the hazard function with the shape of the λ_o being dictated by the assumed distribution.

$$\lambda_i(t, \mathbf{x}_i) = \lambda_0 \exp\{\mathbf{x}_i{}'\beta\} \tag{7}$$

For the exponential distribution, the baseline hazard, λ_0 is constant. The hazard rates for various covariates $\mathbf{x}_i$ with their statistical significance are displayed in Table 17. In the following set of tables for all the analysis results, the hazard rates will be used. Since the hazard rates are the division of the effect of a certain covariate by a baseline hazard, a hazard rate lower than 1 can be interpreted as meaning that the covariate lowers the likelihood of the event (exit) occurring, and greater than 1 will imply that the covariate increases the likelihood of the event occurring. Clearly, if it is equal to 1, then it has no effect.

Secondly, the hazard rates of the continuous covariates have been transformed into a more easily interpretable form in Table 5 with the equation:

$$\exp(\beta\, log(1.01)) = 1.01^{\lambda_i}$$

where, the hazard ratio $\lambda_i = \exp\beta$, and the 1.01 is the 1% increase in the selected covariate. We have presented the hazard rate in the following tables.

Next, we will present the tables for the Weibull Survival estimate in Table 8 and Table 9 results for the other models are displayed in the appendix.

Table 8: Weibull survival coefficients for the ISO-500 list

	(1)	(2)	(3)	(4)
Log of the Number of Employees Before the Exit Year	0.708*** (-10.26)	0.571*** (-9.98)	0.703*** (-9.41)	0.578*** (-8.94)
Secular Network	0.973 (-0.19)	0.937 (-0.37)	1.061 (0.36)	1.042 (0.19)
No Network	2.058*** (5.60)	1.783*** (3.66)	2.031*** (4.74)	1.965*** (3.47)
Log Entry Productivity	0.918** (-2.65)	0.998 (-0.04)	0.939 (-1.78)	1.062 (1.06)
Log Entry Profitability	1.271*** (16.81)	1.249*** (12.58)	1.278*** (16.45)	1.264*** (12.24)
Log Exit Productivity	0.827*** (-6.12)	0.804*** (-4.49)	0.849*** (-4.83)	0.825*** (-3.60)
Log Exit Profitability	0.942** (-2.74)	1.011 (0.34)	0.927*** (-3.34)	1.029 (0.81)
Export	0.523*** (-7.95)	0.502*** (-4.27)	0.593*** (-5.71)	0.635* (-2.49)
Alpha	0.276*** (11.78)	0.363*** (11.20)	0.352*** (15.01)	0.436*** (13.40)
Significance of the Constant Industry Dummy City Dummy	*** No No	*** Yes No	*** No Yes	*** Yes Yes
N	1140	658	1140	658

Exponentiated coefficients; t statistics in parentheses
="* p<0.05 ** p<0.01 *** p<0.001"

Table 9: Weibull survival coefficients for the ISO-1000 list

	(1)	(2)	(3)	(4)
Log of the Number of Employees Before the Exit Year	0.575*** (-14.03)	0.507*** (-13.40)	0.546*** (-13.86)	0.502*** (-12.53)
Secular Network	1.070 (0.49)	1.140 (0.82)	1.057 (0.37)	1.109 (0.59)
Religious Network	1.797*** (5.13)	1.816*** (4.41)	1.665*** (3.98)	1.668*** (3.52)
Log Entry Productivity	0.732*** (-8.27)	0.790*** (-5.06)	0.668*** (-9.22)	0.701*** (-6.24)
Log Entry Profitability	1.449*** (14.91)	1.441*** (11.87)	1.497*** (15.07)	1.484*** (11.84)
Log Exit Productivity	0.805*** (-6.23)	0.763*** (-6.04)	0.823*** (-4.57)	0.813*** (-3.72)
Log Exit Profitability	0.931** (-2.72)	0.966 (-1.07)	0.954 (-1.67)	0.951 (-1.43)
Export	0.883 (-1.23)	0.801 (-1.59)	0.871 (-1.25)	0.932 (-0.47)
Alpha	0.331*** (12.64)	0.420*** (14.36)	0.416*** (15.69)	0.470*** (16.13)
Significance of the Constant Industry Dummy City Dummy	*** No No	*** Yes No	*** No Yes	*** Yes Yes
N	1044	845	1044	845

Exponentiated coefficients; t statistics in parentheses
="* p<0.05 ** p<0.01 *** p<0.001"

The results for this model (similar to the other models) are as expected. The number of employees before the final year has a positive impact on survival (the higher the number of employees, the lower the hazard for exit), entry and exit productivity, as well as exit profitability are also positively correlated with survival, along with the exporter status.

Looking at the network status of the firms, it is clearly visible that membership in a network is positively correlated with survival, as non-networked firms have 1.97 times the hazard rate of a religious firm in the ISO-500 list and. By the Weibull estimate, the secular networks do not appear to have statistically significant difference to religious networks. In fact, this lack of a difference is repeated across all models (see Appendix), with secular and religious networks performing very similarly in the ISO-500 list. One such reason for this similarity can be seen in Figure 22, where the survival curves for the secular and religious networked firms were within each other's 95% confidence interval in the ISO-500 list. These results are similar to the ISO-1000 list.

Comparatively, in the Weibull distribution, the effect of the number of employees seem to be higher than the Exponential Model (see Table 5 as well as the Appendix for the comparisons) (0.55% decrease in hazard for 1% for the ISO-500 list, and 0.68% decrease in hazard for 1% for the ISO-1000 list), as well as the effects of secular and religious networks, with a firm in a secular network being 53% less likely to experience exit, and a religious network 51% less likely. However, their significance remains the same, as well as their effect, meaning higher number of employees and belonging to a network increase firm survival. Exit productivity is still relevant to the firm survival as well, with an increase of 1% reducing the hazard by 0.19%. In the ISO-1000 list, exporter status loses significance once both the industry and city dummies have been accounted for but is still positively correlated with firm survival.

The alpha value is between 0 and 1, and this implies that the hazard decreases with duration. The non-parametric estimation had a shape of the cumulative hazard curve which was concave, implying that hazards decreased with time, and that observation is corroborated by the alpha value. As firms stay in the list, they are more likely to remain there. Whether this decrease in hazard monotonic, or first increases before decreasing for the longest-surviving firms cannot be deduced from this model, but the non-parametric graph of the integrated hazard appears monotonically concave, and therefore, in the log-logistic model, we expect to find a monotonic decrease in hazard. This, too, is the case, as the Log-logistic model's coefficients are between 0 and 1, implying a monotonic decrease in hazard. (See the Appendix for the table of covariates as well as the analysis of the Log-Logistic model)

For the rest of the covariates, the functions with the other distributions and the Weibull distribution behave similarly. Neither entry productivity nor the exit profitability has any significant effect on the survival of the firm, and entry profitability still appears to be negatively correlated with firm survival (0.23% and 0.39% increase in hazard per 1% increase in the ISO-500 and ISO-1000 lists respectively).

ISO-500 and ISO-1000 behave similarly, except for the exporter status, which has no significance for the ISO-1000 list. As with the exponential distribution, this may be due to the proportion of exporters in both lists.

We will next present a comparison of the survivor functions fitted across different models that we have used for the ISO-500 and ISO-1000 lists.

Figure 23: Comparison of the different estimates for the ISO-500 list

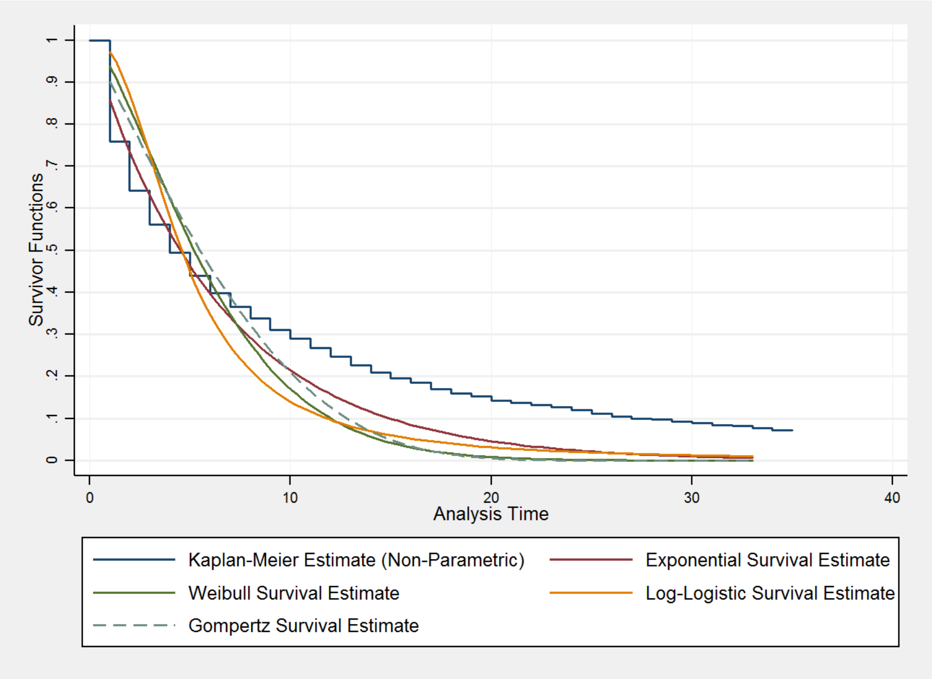
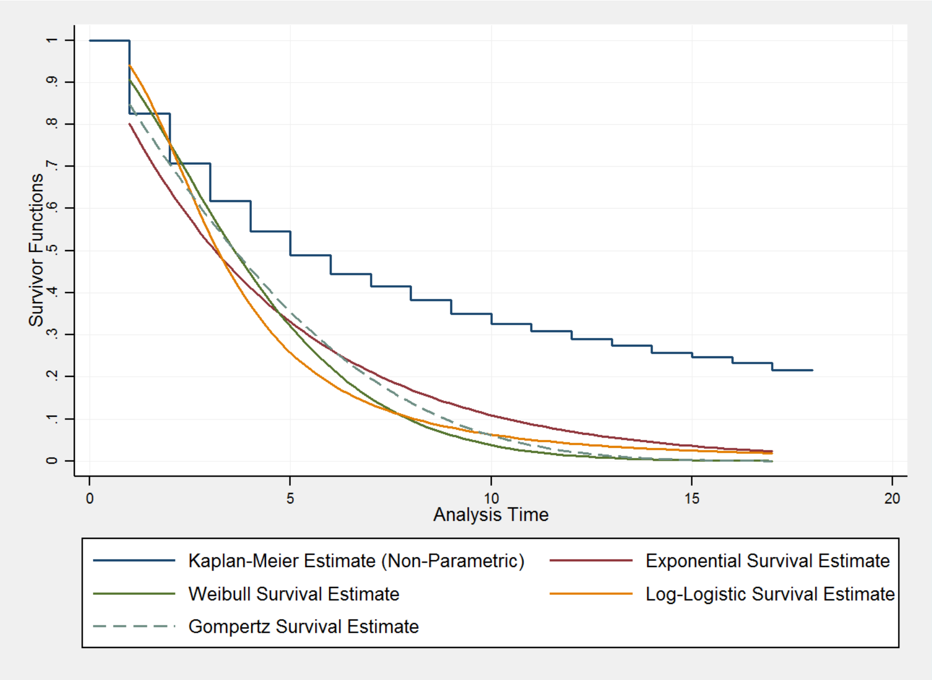


Figure 24: Comparison of the different estimates for the ISO-1000 list



6.3. The Entry Profitability Issue

While a positive correlation with entry profitability may be seen either as a fluke or an artefact of the model, it appears too consistently to not require further analysis.

One possible reason for the negative correlation with survival may be the high rate of 1-year entries. Therefore, one possible way to gauge the issue would be by removing the 1-year entries and running the analysis. This too was done across models, and the results for the exponential model for the ISO-500 list are presented in Table 10. As the other models also do not differ greatly, they have been omitted.

The results for all the hazard ratios are similar, and there does not appear to be any significant change in the effect of entry profitability.

Table 10: Exponential survival function for the ISO-500 list omitting the 1-year entries

	(1)	(2)	(3)	(4)
Log of the Number of Employees Before the Exit Year	0.756*** (-6.87)	0.664*** (-6.57)	0.763*** (-6.07)	0.672*** (-5.78)
Secular	0.566*** (-5.71)	0.603*** (-3.99)	0.633*** (-4.25)	0.676** (-2.81)
Religious	0.559*** (-4.03)	0.637** (-2.65)	0.610** (-2.96)	0.674 (-1.93)
Log entry Productivity	0.939 (-1.62)	0.973 (-0.51)	0.958 (-1.01)	1.015 (0.24)
Log Entry Profitability	1.161*** (10.09)	1.148*** (7.60)	1.156*** (9.22)	1.145*** (6.86)
Log Exit Productivity	0.862*** (-4.01)	0.834*** (-3.45)	0.874*** (-3.42)	0.851** (-2.83)
Log Exit Profitability	0.991 (-0.37)	1.014 (0.38)	0.979 (-0.79)	1.020 (0.53)
Export	0.742** (-2.76)	0.649* (-2.31)	0.827 (-1.57)	0.816 (-0.96)
N	823	534	823	534

Exponentiated coefficients; t statistics in parentheses
="* p<0.05 ** p<0.01 *** p<0.001"

As both the profits and the revenue (as well as the total assets) of the firms are adjusted for inflation, any issue arising from the changes to the money are also discounted.

Another possible solution to this is to accept that the first-year profitability is indeed negatively correlated with survival: but why? A speculative answer to that would be that as the ISO lists are compiled over sales, a firm may be incentivized to maximize sales in order to one of the top 500/1000 firms. Therefore, firms may make inefficient investments which would yield an apparent high return in the first year (i.e. the entry year into the list) but may fall off over time, whereas, a firm which entered the list without making these decisions would enter with a lower profitability, which would rise over time. Therefore, to test this, we have also looked at the motion of the profitability of firms over time for both firms that have remained in the list for a short duration, and compared them with more resilient firms. As the mean survival duration in the ISO-1000 list is around 6.25 years, for our analysis we had two groups: firms which were on the list for less than 7 years or firms which were in the list for 7 years or more. Regressing the log of their profitability on the `run` variable (i.e. the number of years the first is on the list), gave us Table 11:

Table 11: Log of profitability for firms that survived for long and short durations

	(1)	(2)
Ln(Prof) if total time in list >= 7 years	0.104** (2.59)	
Ln(Prof) if total time in list < 7 years		0.0213 (1.71)
Constant	9.819*** (72.66)	2.932*** (68.01)
N	6262	8872
t statistics in parentheses ="* p<0.05	** p<0.01	*** p<0.001"

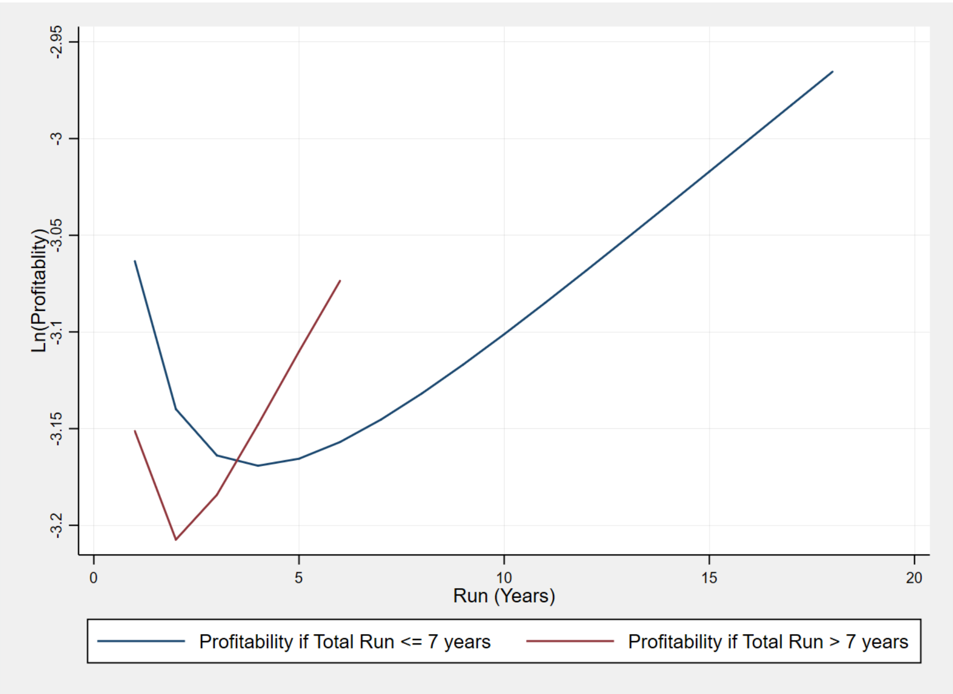
As can be seen, the profitability is positively correlated with run years in both but only the long surviving firms have a statistically significant correlation. This does not explain, however, the negative correlation between first year profitability and survival. When this regression is done only for the first three years of a firm’s run, however, the picture is quite different:

Table 12: Log of profitability for firms that survived for long and short durations for the first three years of survival

	(1)	(2)
Ln(Prof) if total time in list >= 7 years	-0.0116 (-0.50)	
Ln(Prof) if total time in list < 7 years		-0.00457 (-0.61)
Constant	1.855*** (23.37)	1.797*** (68.92)
N	686	5834
t statistics in parentheses ="* p<0.05	** p<0.01	*** p<0.001"

Here it is clearly visible there is a negative correlation between time spent in the list and profitability: profitability decreases for the first three years. This can also be seen in the graph of the fit lines.

Figure 25: Fit lines for profitability across years for a a firm is in the ISO-1000 list



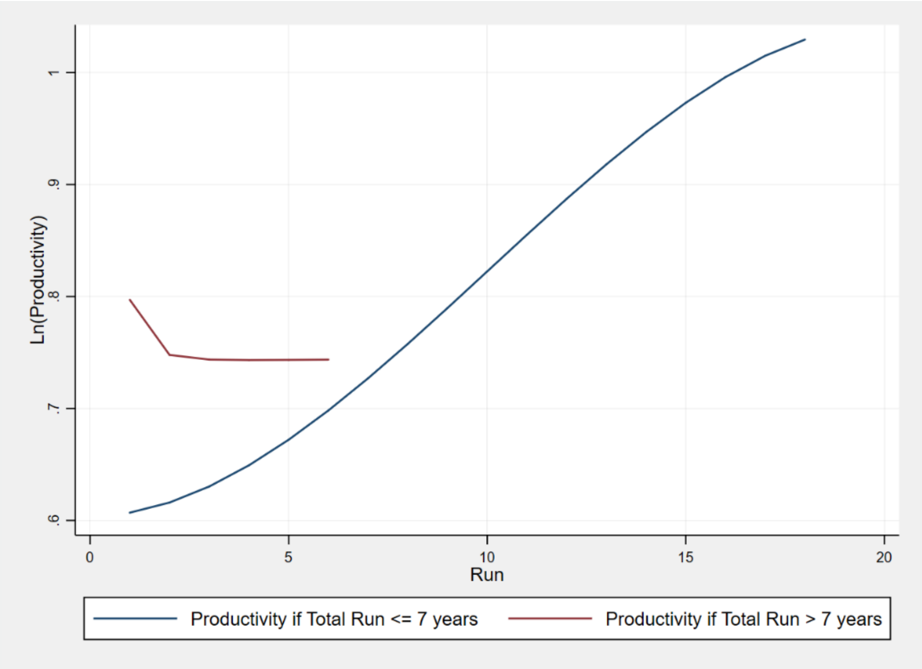
This clearly shows that the profitability of the firms take a dive for the first three years that a firm is in the list and then increase, with the net effect being positive.

There are two obvious possibilities why this may be the case: (1) there is an issue arising from the calculation of the PPE, (2) this is only true because the first cohort was in 1997, and there is a disproportionately high number of failures in the economic crises of 1999 and 2001, skewing the results. Both hypotheses can be indirectly tested.

For (1), we can look at another metric that makes use of the PPE – namely Productivity. We can see the motion of the Productivity in Figure 26.

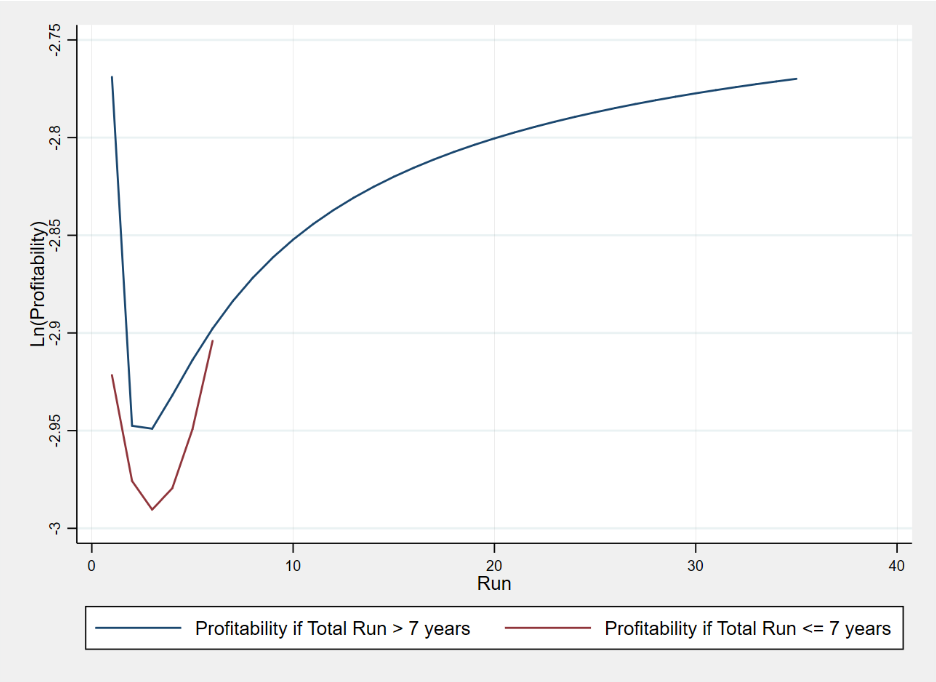
In that figure we can clearly see that the productivity for long-term surviving firms increase across time, whereas the short-term surviving firms start higher and then take a dip. In fact, on average, the productivity of short-term surviving firms starts higher than those of long-term surviving firms. This agrees with our initial hypothesis that short-term surviving firms may have illusory productivity by making bad investments (thus higher Value Added in the short term). However, as the Productivity of the long-term surviving firms keep increasing without a dip, we can probably rule out an issue arising from the PPE.

Figure 26: Fit lines for productivity across years a firm is in the ISO-1000 list



The second possible issue is that this is a case unique to the ISO-1000 list, on account of specific circumstances relating to its starting cohort. We can also test this by looking at the motion of profitability across years for the ISO-500 list, which began in 1980.

Figure 27: Fit lines for productivity across years a firm is in the ISO-500 List



However, the trend observed for the ISO-1000 list appears to be even amplified for the ISO-500 list. Therefore, the entry cohort does not appear to be the issue either.

Since, this pattern repeats across the lists, we need to ask which of the values (or possibly both) used to calculate the profitability behave in this way, to produce these results. As discussed in the data section, profitability is calculated as the Profits/Revenue. It was also calculated as Profits/Total Assets, and in this specification, the behavior remained. Therefore, an initial guess would be that the issue stems from Profits. Indeed, looking at the Real Profits for the ISO-1000 list, we can see a dip after the first year run, which would effect the profitability. This is almost certainly due to the high number of firms that were on the list for one year only. There is an issue with this, however. As we have stated above, we have run the survival analysis excluding firms that were on the list for more than one year (Table 10) and the negative correlation between first year profitability and survival remained. Also, looking at Figure 25 we can clearly see that the dip in profitability is not only one year but continues for the next few years as well. Further, the issue with the one-year firms' profits should also extend to other micro data such as Real Value Added or Real Revenue– which it does not. (Figure 29 for Real Value Added, Real Revenue is omitted as it is similar in shape to the Real Value Added)

Figure 28: Real profits of firms across run years (ISO-1000)

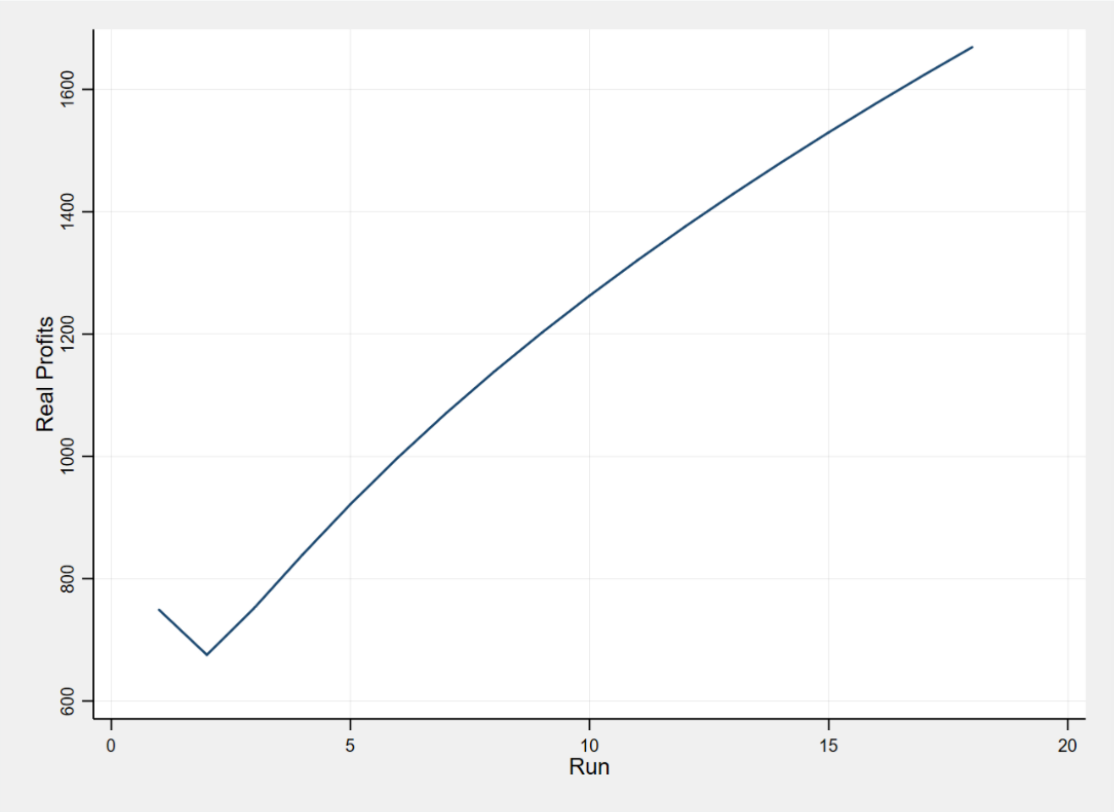
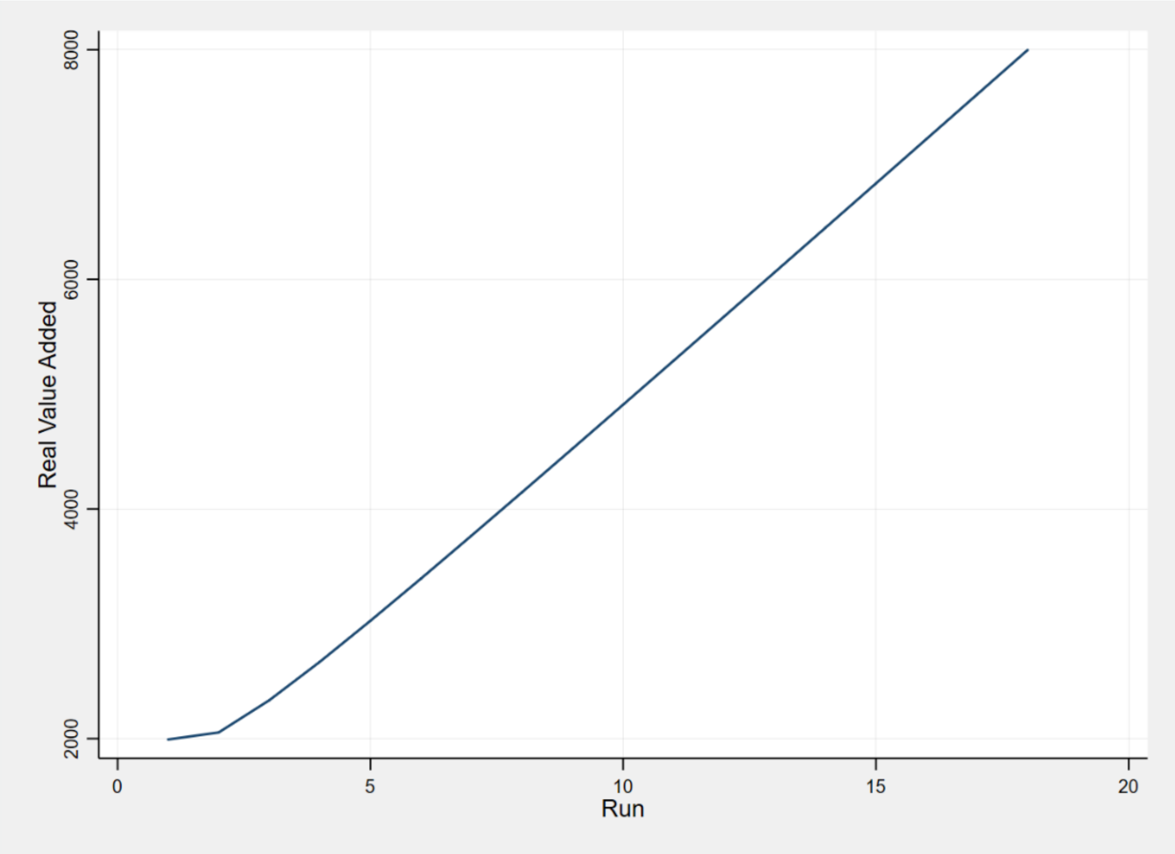


Figure 29: Real value added of firms across run years (ISO-1000)



Therefore, while there are some possible explanations, overall, we do not have a fully convincing answer to this issue and it remains an open problem. We can, conclude, however, that at least for productivity, our initial hypothesis of quick jump in productivity and then slowly falling off is supported.

6.4. Survival Analysis with Time Varying Covariates

While the survival analysis has given us certain insights into the resilience of the firms given time invariant covariates such as the firm productivity and profitability in its entry and exit years, the list gives us more data to work with. Before giving the hazard rates, we will again present the effects of 1% increase in the continuous covariates for a baseline model of ln(Run) (although the results are similar across models) in Table 13:

Table 13: Percent changes in the survival with 1% increase in the continuous covariates for the ln(Run) model

	ISO-500 (%)	**ISO-1000 (%)**
1% increase in:		
Number of Employees	-0.94	-0.78
Productivity	-0.76	-0.93
Profitability	-0.93	-0.89

All above results are statistically significant. For 1% increase in the number of employees, the hazard rates fall by 0.94%, for productivity, by 0.76%, and for profitability by 0.93% in the ISO-500 list, by 0.78%, 0.93% and 0.89% in the ISO-1000 list respectively.

As we have the firm's employee size, revenue and profits for each year that it is in the list, we are able to look at the survival of the firms given these yearly varying inputs. For the ISO-500 and ISO-1000 lists, we have the following rates in Table 14 and Table 15 .

Table 14: Time varying complementary Log-Log model for the ISO-500 list

	(1)	(2)	(3)	(4)
Secular	1.284* (2.03)	1.321* (2.24)	1.327* (2.27)	0.949 (-0.43)
No Network	2.555*** (8.67)	2.736*** (9.14)	2.780*** (9.27)	1.888*** (5.73)
Log Size	0.851*** (-7.56)	0.817*** (-9.55)	0.814*** (-9.77)	0.715*** (-11.11)
Export Dummy	0.606*** (-6.52)	0.593*** (-6.74)	0.592*** (-6.75)	0.562*** (-7.56)
Log Productivity	0.897*** (-4.32)	0.879*** (-5.17)	0.877*** (-5.29)	0.839*** (-7.04)
Log Profitability	0.967* (-2.01)	0.965* (-2.13)	0.964* (-2.18)	1.007 (0.40)
Baseline Hazard	0.786*** (-8.25)	1.000* (-2.34)	1.000 (-1.62)	
Functional Form of Baseline Hazard	ln(Run)	Run^2	Run^3	6 discrete period dummies
N	7506	7506	7506	7506
Exponentiated coefficients; t statistics in parentheses ="* p<0.05	** p<0.01	*** p<0.001"		

In this regression, as well as the next one, columns (1)-(4) all include the city and industry dummies but have different functional forms for the baseline hazard: ‘Baseline Hazard’ has been defined as *ln(run)* in column (1), *run*^{2} in column (2), run^3 in column (3), and as a set of discrete duration values in column (4). Duration 1 takes on 1 for the first 6 years of a firm’s run, Duration 2 takes on 1 for 7-12 years, etc. For a firm that has had a 14-year run in the list, Duration 1-3 would be 1, and Duration 4-6 would be 0.

Looking at column (4), the baseline hazard for belonging to a secular network is not statistically different than belonging to a religious network (although slightly lower) and being unaffiliated increase the hazard rate by 2-fold.

This is particularly close to the time-invariant functions that had the Weibull or log-logistic distributions (in Table *8* and the appendix respectively) with hazard functions lowered by around 0.5 in the Weibull distribution and by 0.3 in the log-logistic. However, before looking at the effects of productivity and profitability, as well as that of firm size, we will add the results for the ISO-1000 list in Table 15.

Table 15: Time varying complementary Log-Log model for the ISO-1000 list

	(1)	(2)	(3)	(4)
Secular	1.173 (1.42)	1.172 (1.41)	1.172 (1.41)	0.921 (-0.76)
Religious	2.465*** (9.78)	2.482*** (9.84)	2.497*** (9.91)	1.813*** (6.47)
Log Size	0.777*** (-11.24)	0.769*** (-11.89)	0.766*** (-12.13)	0.603*** (-16.22)
Export Dummy	0.772** (-2.60)	0.767** (-2.67)	0.767** (-2.66)	0.639*** (-4.85)
Log Productivity	0.830*** (-6.99)	0.831*** (-6.96)	0.831*** (-6.94)	0.760*** (-10.81)
Log Profitability	0.930*** (-4.15)	0.932*** (-4.08)	0.931*** (-4.10)	1.015 (0.74)
Baseline Hazard	0.908** (-2.95)	0.999 (-1.90)	1.000 (-1.46)	
Functional Form of Baseline Hazard	ln(Run)	Run^2	Run^3	3 discrete period dummies
N	14817	14817	14817	14817

Exponentiated coefficients; t statistics in parentheses
="* p<0.05 ** p<0.01 *** p<0.001"

The behavior of the firms in secular networks or non-networked do not significantly differ from the ISO-500 list, with secular networks performing slightly better (not statistically significant) than religious networks, and non-networked firms having around 1.8 times the hazard ration of the religious networked firms. Further, the hazard of the size of the firm (i.e. the number of employees) is 0.39 in the ISO-500 list, and around 0.51 for the ISO-1000 list) are similar to hazards for the final year of the firm in the time-invariant models. This may mean that the effect of size on exit is determined solely by the effects of the size in the final year. Similarly, the effects of the productivity are similar to the exit productivity (although, in the ISO-1000 list, the entry productivity also had a significant effect).

One note to make here is that unlike the time-invariant model which could only capture the entry and exit profitability of firms and appeared to produce a counterintuitive negative correlation with survival with entry profitability, across time lists consistently show that profitability is positively correlated with firm survival.

In the following pages, Figure 30 and Figure 31 show the combined Survival, Hazard and Cumulative Hazard rates for the ISO-500 and ISO-1000 rates respectively. We should note that we can clearly see that the hazard rates fall across run, again confirming our non-parametric observation that cumulative hazard appears to be concave. Another interesting note is that the cumulative hazard for the ISO-500 seems to be increase sharper in the first 10 years before tapering down, whereas the ISO-1000 list's cumulative hazard increases more gradually. We can again reference Figure 20 and Figure 21 to see that the firm size in sales decreases exponentially in the first 500 but much more linearly in the next 500, which may explain the relative instability of the smaller list.

Figure 32 And Figure 33.show the Survival rates separated for firms in religious and secular networks as well as unaffiliated ones. This again confirms both our initial observations and the time-invariant analysis, and therefore, there is no need to further analyze them.

Figure 34 and Figure 35 show the Hazard rates by network by the list. Here we can see that while the hazard rates fall across time that a firm has been in the list, the initial hazard rates for firms in networks (secular more than religious) start low, but all converge in slope as a firm's time in the list increases, with the hazards for unaffiliated firms falling sharply in the first 10 years in the ISO-500 list, more gradually in the ISO-1000 list. One interesting thing to note here is that in the ISO-1000, it appears that the hazard rates converge for religious and secular networks, and in fact, in the final years, religious networks appear to have less hazard. This is not repeated in the ISO-500 list. We have also looked at whether this time also translates to the years that the firm is active in, as the political situation after 2002 may have favored religious networks over secular ones, but year by year this convergence cannot be observed, with secular firms consistently having fewer hazards to survival.

Finally, Figure 36 and Figure 37 compare both the size and the shape of the graphs for the two lists. Again, unsurprisingly, ISO-1000 list has a higher survival rate and lower hazard rates than the ISO-500 list.

Figure 30: Fitted hazards, integrated hazards, and survival across time for ISO-500

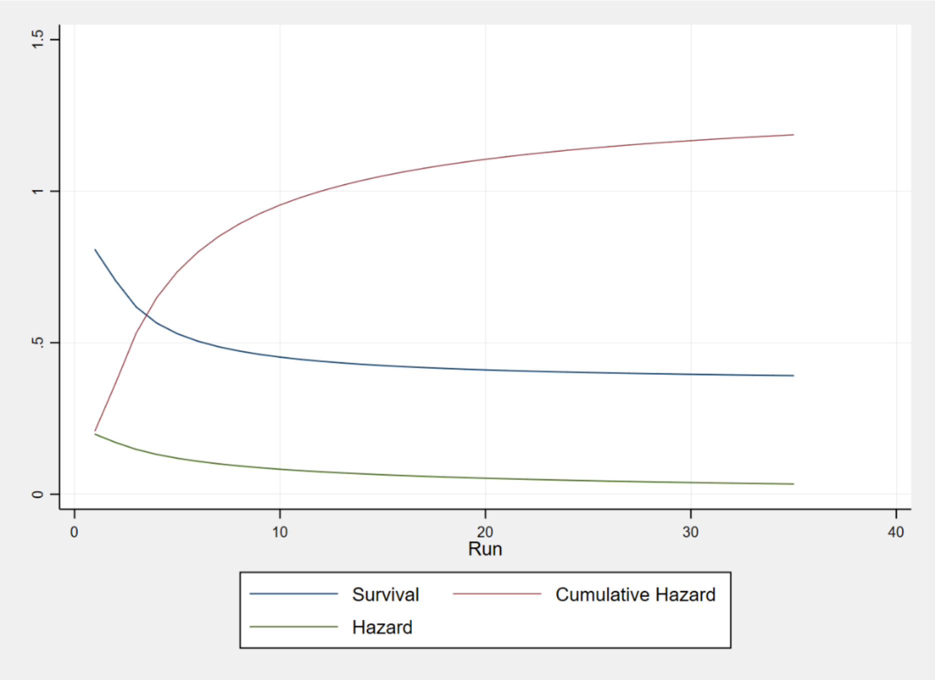


Figure 31: Fitted hazards, integrated hazards, and survival across time for ISO-1000

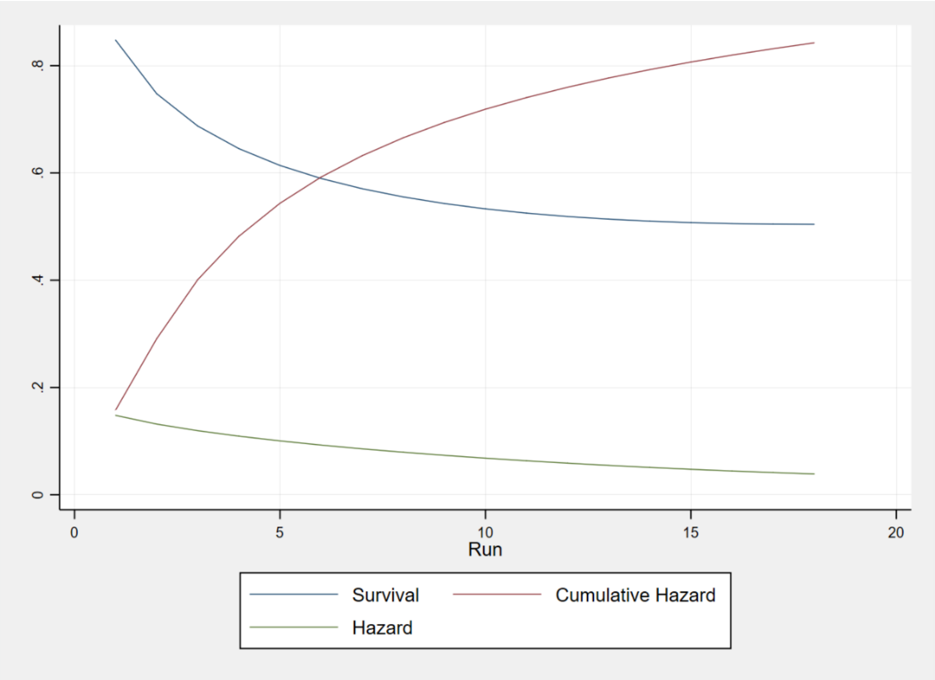


Figure 32: Comparison of fitted survival across time by network (ISO-500)

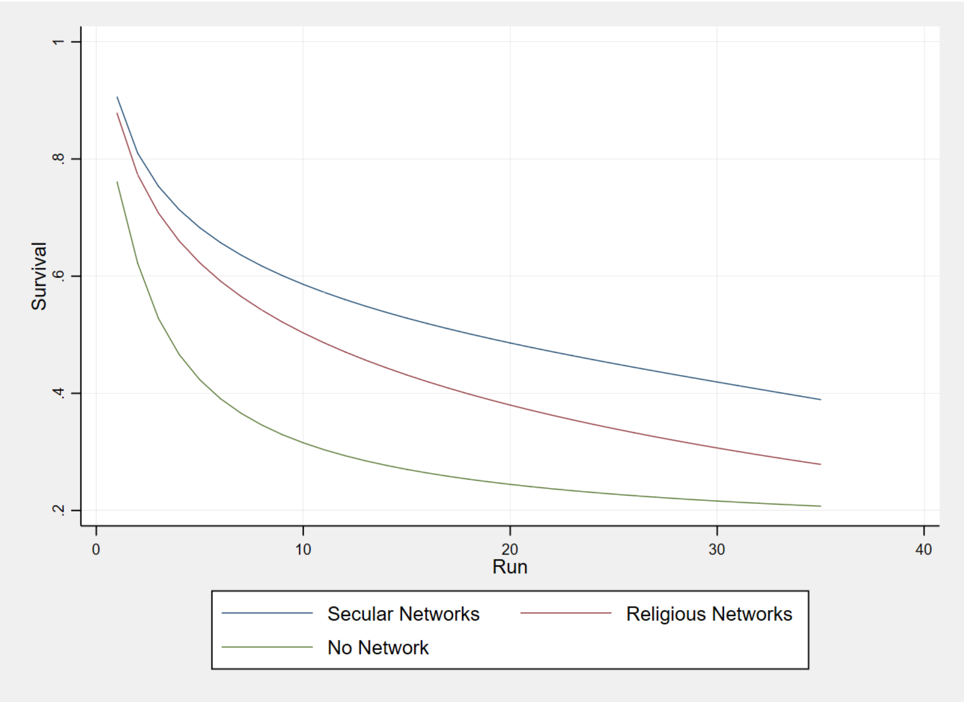


Figure 33: Comparison of fitted survival across time by network (ISO-1000)

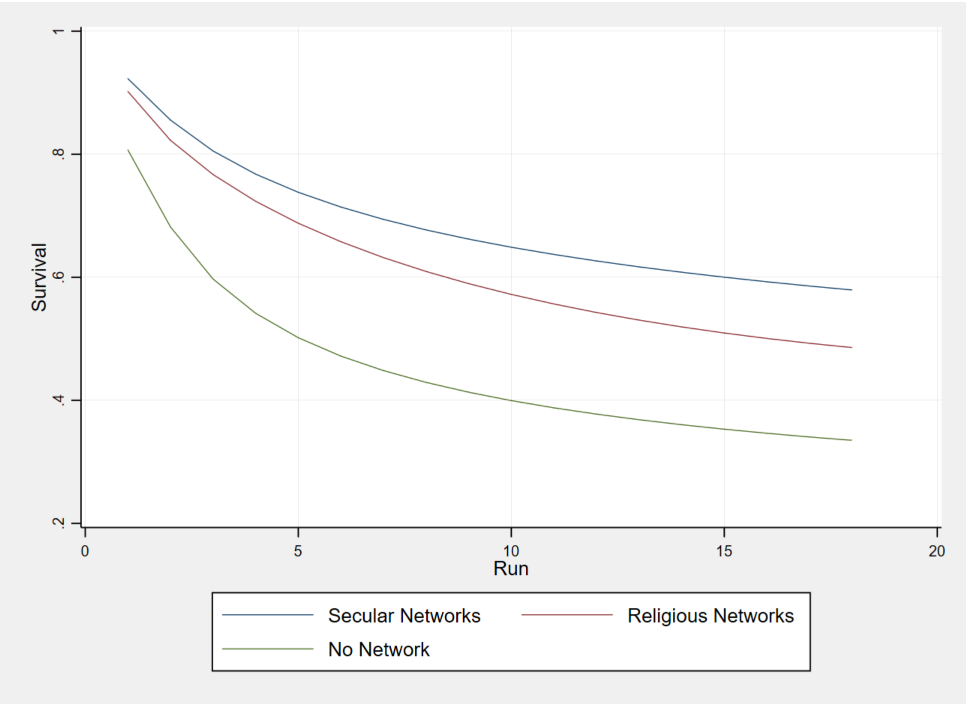


Figure 34: Fitted hazards across time by network (ISO-500)

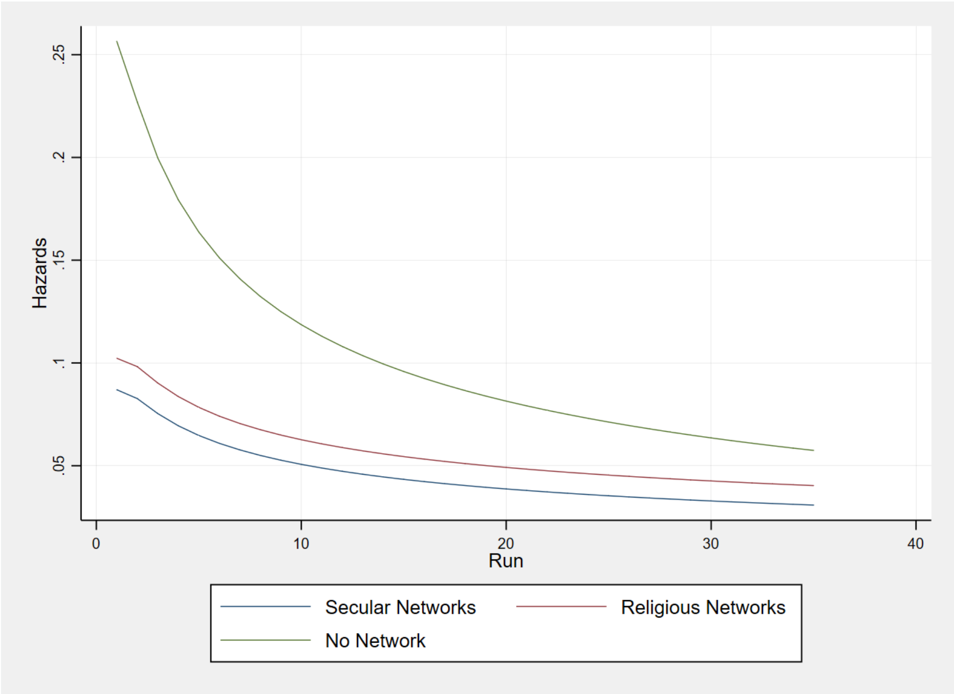


Figure 35: Fitted hazards across time by network (ISO-1000)

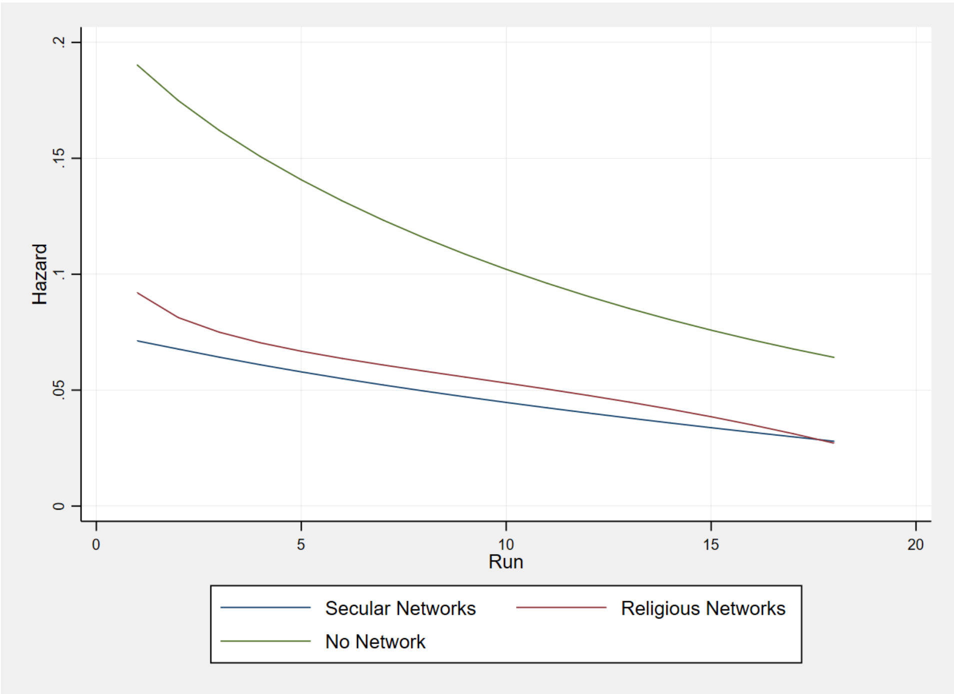


Figure 36: Comparison of survival functions between the ISO-500 and ISO-1000 lists

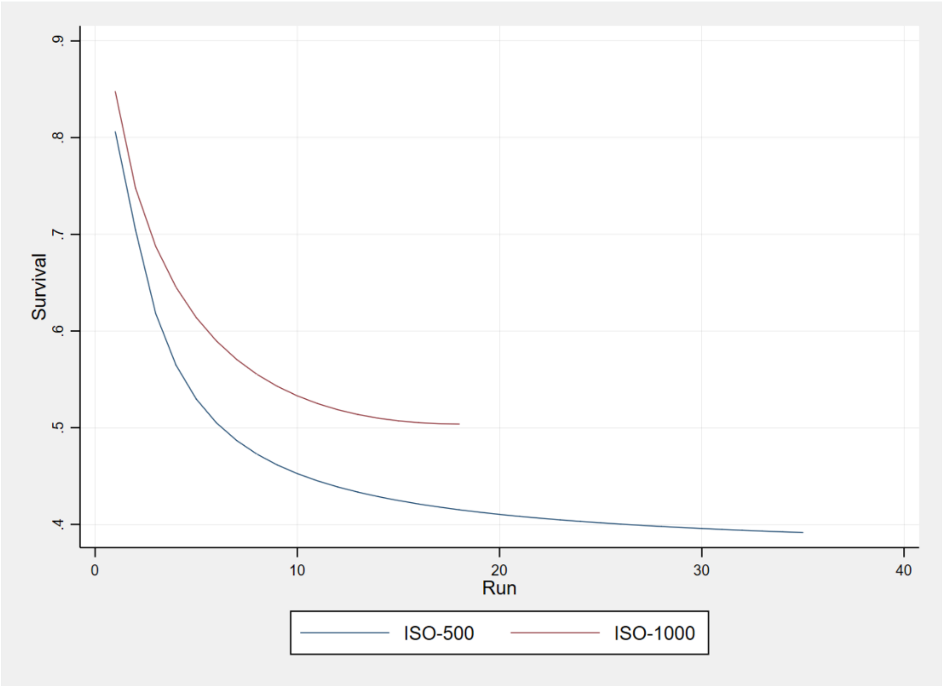


Figure 37: Comparison of hazard functions between ISO-500 and ISO-1000 lists

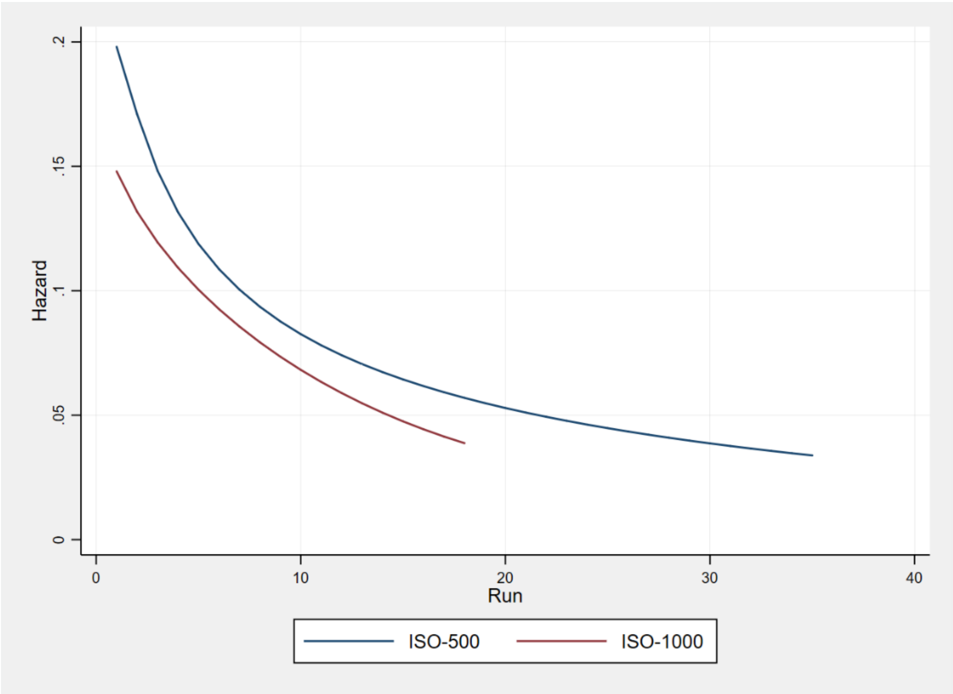


Table 16: Hazard rate comparison of the different models

	ISO-1000						ISO-500					
	Time Variant	*Cox*	*Exp.*	*Weibull*	*Log-Logistic*	*Gompertz*	*Time Variant*	*Cox*	*Exp.*	*Weibull*	*Log-Logistic*	*Gompertz*
Secular	0.508***	0.736*	0.761*	0.665***	0.180*	0.700**	0.385***	0.563***	0.653***	0.530***	0.269***	0.594***
	(-9.06)	(-2.53)	(-2.27)	(-3.31)	-2.3	(-2.90)	(-8.63)	(-4.37)	(-3.41)	(-4.80)	-3.39	(-8.61)
Religious	0.550***	0.659**	0.684**	0.600***	0.308***	0.622***	0.395***	0.562**	0.626*	0.509***	0.294*	0.513***
	(-6.48)	(-2.92)	(-2.67)	(-3.52)	-3.58	(-3.29)	(-5.34)	(-3.02)	(-2.53)	(-3.47)	-2.48	(-4.99)
Log Size	0.603***	0.582***	0.610***	0.502***	0.405***	0.537***	0.699***	0.622***	0.668***	0.578***	0.340***	0.517***
	(-16.22)	(-9.98)	(-9.53)	(-12.53)	-12.09	(-11.37)	(-8.51)	(-7.87)	(-6.88)	(-8.94)	-8.55	(-3.42)
Export Dummy	0.639***	0.972	0.982	0.932	0.0369	0.127***	0.518***	0.744	0.787	0.635*	0.265*	0.699*
	(-4.85)	(-0.20)	(-0.13)	(-0.47)	-0.41	-8.7	(-6.46)	(-1.66)	(-1.36)	(-2.49)	-2.37	(-2.01)
Log Productivity	0.760***	0.779***	0.804***	0.701***	0.178***	0.746***	0.792***	1.036	1.024	1.062	-0.0027	1.055
(Log Entry Prod.)	(-10.81)	(-4.43)	(-3.94)	(-6.24)	-4.7	(-5.18)	(-7.13)	(0.63)	-0.42	-1.06	(-0.07)	-0.92
(Log Exit Prod.)		0.840**	0.847**	0.813***	0.101**	0.816***		0.842**	0.870**	0.825***	0.0884*	0.826***
		(-3.13)	(-3.00)	(-3.72)	-2.66	(-3.64)		(-3.19)	(-2.62)	(-3.60)	-2.49	(-3.55)
Log Profitability	1.015						1.008					
	(0.74)						-0.31					
Industry Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Exponentiated coefficients; t statistics in parentheses

="* p<0.05 ** p<0.01 *** p<0.001"

THIS PAGE IS INTENTIONALLY LEFT BLANK

7. CONCLUSIONS AND FURTHER RESEARCH

While we have previously given our temporary conclusions piecemeal, we would both like to reiterate our findings and also see whether our hypotheses hold.

We can refer to Table 16, we have made the comparison between the different models, and from all the models, we can make a the following conclusions regarding our hypotheses:

1) For productivity and profitability, we can conclude that:
    a. Higher the productivity of the firm, especially in its final year, the lower is its hazard for exit.
    b. Profitability also increases firm survival, although first year profitability appear to have a negative impact, which according to our model remains unexplained.
2) Most of the models agree that being an exporter increases the chances of survival, but the specific significance of exporter status changes with the model.
3) All models unambiguously agree that higher the size of the firm, especially in its final year, the lower is its hazard for exit.
4) Firms belonging to networks have lower hazards than unaffiliated firms, with firms belonging to secular networks having higher survival rates across all models than those belonging to religious networks.
5) All models agree that operating out of major industrial centers effects the survival positively, with Ankara, Bursa, Gaziantep, Istanbul, Izmir, Kayseri appearing to be positively correlated with survival both across models, and across the ISO-500 and ISO-1000 lists.
6) The effects of industry is a little more complicated but we can conclude that among the top 500 firms belonging to mid-technology sectors such as industrial chemicals, metals, metal products consistently appear to be related with lower hazards across all models, but among the top 1000 (and therefore the second 500), other mid-technology industries

such as iron products, petroleum, plastics and scientific equipment low-technology industries like apparel, beverages, paper and textiles appear to be positively correlated with survival.

We have pointed out before that we did not find a convincing reason why entry profitability is negatively correlated with firm survival, and therefore, a further area of research would be look at that relation to try and explain it. When we used the profitability of a firm across the years (not presented in this paper) it also is positively related with survival. Also, when almost all of the data has been used – i.e. using the full data with the analysis on the varying covariates – this counterintuitive relation disappears. Therefore, while we do not believe this result to be particularly relevant but as we cannot explain it and it remains across different models, it might be an interesting area to look.

Also, the ISO lists are a treasure trove of further data, a few of which we did not use for our research. There is further data on the public and private ownership percentages of the firms, foreign ownership percentages of the firms, all of which can also be utilized in the context of firm survival among the top firms.

Similarly, we have looked at the survival of firms among the top in Turkey, but their actual positions as well as their sales figures can also be used for further research.

BIBLIOGRAPHY

Agarwal, Rajshree; Gort, Michael (2002), "Firm and Product Life Cycles and Firm Survival", American Economic Review, Volume 92, No. 2 (May 2002), pp. 184-190

Atiyas, Izak; Bakis Ozan, Gurakar; Esra Ceviker (2016), "Anatolian Tigers and the Emergence of the Devout Bourgeoisie in the Turkish Manufacturing Industry: An Empirical Analysis", Economic Research Forum, Working Paper No. 1064 (2016)

Audretsch, David B. (1991), "New-Firm Survival and the Technological Regime", The Review of Economics and Statistics, Volume 73, No. 3 (1991), pp.441-450

Audretsch David B.; Mahmood, Talat Dr. Dr. (1995), "New Firm Survival: New Results Using a Hazard Function", Review of Economics and Statistics, Volume 77 (February 1995), pp. 97-103

Bruderl, Josef; Preisendorfer, Peter (1992), "Survival Chances of Newly Founded Business Organization", American Sociological Review, Volume 57, No.2 (April 1992), pp. 227-242

Buddelmeyer, Hielke; Jensen, Paul H.; Webster, Elizabeth (2006), "Innovation and the Determinants of Firm Survival", Institute for the Study of Labor, Discussion Paper, No. 2386

Cefis, Elena; Marsili, Orietta (2005), "A Matter of Life and Death: Innovation and Firm Survival", Industrial and Corporate Change, Volume 14, No. 6 (2005), pp. 1167-1192

Ciuca, Vasilica; Matei, Monica Mihaela Maer (2010), "Survival Analysis for the Unemployment Duration", Proceedings of the 5th WSEAS International Conference of Economy and Management Transformation, Volume I (January 2010)

Cox, D.R (1972), "Regression Models and Life-Tables", Journal of Royal Statistical Society Series B (Methodological), Volume 34, No. 2 (1972), pp.187-220

Crowley, John; Hu, Marie (1977), “Covariance of Heart Transplant Survival Data”, Journal of the American Statistical Association, Volume 72, No. 357 (March 1977), pp. 27-36

Esteve-Perez; Llopis, Amparo Sanchis; Llopis Juan A. Sanchis (2010), “A Competing Risk Analysis of Exit for Spanish Manufacturing Firms”, Empirical Economics, Volume 38, No. 2 (April 2010), pp. 281-304

Esteve-Perez, Silviano; Manez-Castillejo, Juan A.; Rochina-Barrachina, Maria E.; Sanchis-Llopis, Juan A. (2007), “A Survival Analysis of Manufacturing Firms in Export Markets”, Entrepreneurship, Industrial Location and Economic Growth

Holmes, Phil; Hunt, Andrew John; Stone, Ian (2009), “An Analysis of New Firm Survival Using a Hazard Function”, Applied Economics, Taylor & Francis (Routledge), Volume 42, No. 02 (2009), pp.185-195

Jenkins, Stephen P. (2001), “Discrete Time Proportional Hazards Regression”, Stata Technical Bulletin, Volume 39 (2001), pp. 22-32

Jenkins, Stephen P. (2004). Survival Analysis. Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK. Downloadable from http://www.iser.essex.ac.uk/teaching/degree/stephenj/ec968/pdfs/ec968lnotesv6.pdf

Kaplan E.L.; Meier Paul (1958), “Nonparametric Estimation from Incomplete Observations”, Journal of the American Statistical Association, Volume 53, No. 282 (June 1958), pp. 457-481

Kiefer, Nicholas M. (1988) “Economic Duration Data and Hazard Functions”, Journal of Economic Literature, Volume 26, No. 2 (June 1988), pp. 656-679

Lobos, Krzysztof; Szewczyk, Miroslawa (2012), “Survival Analysis: A Case Study of Micro and Small Enterprises in Dolnoslaskie and Opolskie Voivodship (Poland)”, Central European Review of Economic Issues, Volume 15, pp.207-216

Love, James H.; Ganotakis, Panagiotis (2013), “Learning by Exporting: Lessons from high-technology SMEs”, International Business Review, Volume 22, Issue 1 (February 2013), pp.1-17

Madeira, S.; Infante, P; Didelet, F. (2014), “Using statistical models in industrial equipment”, Proceedings of the EuroMaintenance 2014, pp. 451-455.

Pollock, Kenneth H.; Bunck, Christine M.; Winterstein, Scott R.; Chen Chui-Lan (1995), “A Capture-Recapture Survival Analysis Model for Radio-Tagged Animals”, Journal of Applied Statistics, Volume 22, Issue 5-6 (1995), pp. 661-672.

Prentice, R.L.; Gloecker, L.A. (1978), “Regresion Analysis of Grouped Survival Data with Application to Breast Cancer Data”, Biometrics, Volume 34, No. 1 (March 1978), pp. 57-67

Reyes, Eric M.; Thomas, Laine (2014), “Tutorial: Survival Estimate for Cox Regression Models with Time-Varying Coefficients Using SAS and R”, Journal of Statistical Software, No 61, Code Snippet 1

Rodríguez, G. (2007). Lecture Notes on Generalized Linear Models. URL: http://data.princeton.edu/wws509/notes/

Salomon, Robert M,; Shaver, J. Myles (2005), “Learning by Exporting: New Insights from Examining Firm Innovation”, Journal of Economics and Management Strategy, Volume 14, Issue 2 (June 2005), pp. 431-460

Shanklin, William L. (1986), “Fortune 500 Dropouts”, Planning Review Vol. 14 Issue 3 (1986), pp. 12-46

Strangler, Dane; Arbesman Sam (2012), “What Does Fortune 500 Turnover Mean?”, Ewing Marion Kauffman Foundation, June 2012

Vernon, Raymond (1966), “International Investment and International Trade in the Product Cycle”, The Quarterly Journal of Economics, Volume 80, No. 2 (May 1966), pp. 190-207

Yasar, Mahmut; Garcia, Philip; Nelson, Carl H; Rejesus, Roderick M. (2007); “Is There Evidence of Learning-By-Exporting in Turkish Manufacturing Industries”, International Review of Applied Economics, Volume 21, No.2 (2007), pp. 293-305

APPENDIX

Specification for the Cox Proportional Hazards Model

This semi-parametric model, which was first described in Cox (1972) is essentially estimating a separate exponential model at different intervals depending on the baseline hazard estimated non-parametrically. That is to say, the values of λ_0 are estimated non-parametrically (using Partial Likelihoods) and then inserted into the parametric model described by:

$$\lambda_i(t \mid \mathbf{x}_i) = \lambda_0(t) \exp\{\mathbf{x}_i{}'\beta\} \tag{a.1}$$

Let us suppose for a set of firms, $\mathrm{N} = \{1,2 \ldots N\}$, there exist distinct failure times:

$$t_{(1)} < t_{(2)} < t_{(3)} < \cdots < t_{(k)}$$

And during these times, a set of failures occur (i.e. We redefine eq. (4) and (5) as:

$$\hat{\lambda}(t) = \sum_{i=1}^{k} \frac{h_j}{n_j} \delta(t - t_i) \tag{a.2}$$

$$\hat{S}(t) = \prod_{t(\mathrm{i})<\mathrm{t}} \left\{1 - \frac{h_j}{n_j}\right\} \tag{a.3}$$

Where:

- δ is the Dirac Delta function which takes on the value 1 at every event (i.e. failure) within the given time period, $(t - t_i)$ 0 otherwise.
- h_j is the number of completed spells of duration t_j

- n_j is the number of firms that have survived before the completion of t_j (i.e. the set of firms at risk at t_j)

Let us recall Table 3.

Years (j)	h(j)	m(j)	n(j)	λ(j)=h(j)/n(j)	1-λ(j)	S(j)
1	163	462	625	0.2608	0.7392	0.7392
2	119	343	462	0.257576	0.742424	0.5488
3	70	273	343	0.204082	0.795918	0.4368
4	80	193	273	0.29304	0.70696	0.3088
5	86	107	193	0.445596	0.554404	0.1712
6	56	51	107	0.523364	0.476636	0.0816

Here, for duration $j = 5$, (for a failure time at 5 years) the partial likelihood of survival is

$$\hat{\lambda}(5) = 0.445$$

Using this for the baseline hazard function at j=5 we then estimate the hazard and survival functions at piece-wise intervals, the model then approximates a piecewise exponential hazard function. By assuming a different but constant baseline hazard at each time interval, we can also produce a model that would fit the data the best. However, as the results do not necessarily contradict or improve on our other models, we will also present the values in the appendix. This setup of assuming different baseline hazards in different time intervals and then fitting a parametric model at piecewise intervals will also be employed when using the time-dependent covariates.

Results for the Different Models Used in the Time-Invariant Analysis and Their Discussion

Exponential Model

Table 17: Exponential survival hazard ratios for the ISO-500 list

	(1)	(2)	(3)	(4)
Log the Number of Employees Before the Exit Year	0.758*** (-8.51)	0.651*** (-7.96)	0.768*** (-7.30)	0.668*** (-6.88)
Secular	1.048 (0.32)	1.024 (0.13)	1.024 (0.13)	1.043 (0.21)
No Network	1.911*** (5.03)	1.652** (3.21)	1.652** (3.21)	1.598* (2.53)
Log Entry Productivity	0.934* (-2.04)	0.989 (-0.22)	0.948 (-1.44)	1.024 (0.42)
Log Entry Profitability	1.219*** (14.17)	1.189*** (10.04)	1.210*** (13.04)	1.186*** (9.25)
Log Exit Productivity	0.857*** (-4.81)	0.843*** (-3.47)	0.882*** (-3.58)	0.870** (-2.62)
Log Exit Profitability	0.953* (-2.19)	1.001 (0.03)	0.943* (-2.57)	1.009 (0.26)
Export Dummy	0.617*** (-6.09)	0.636** (-2.90)	0.702*** (-3.97)	0.787 (-1.36)
Significance of the Constant Industry Dummy City Dummy	*** No No	*** Yes No	*** No Yes	*** Yes Yes
N	1140	658	1140	658
			XA	aQW
Exponentiated coefficients; t statistics in parentheses ="* p<0.05	** p<0.01	*** p<0.001"		

Here, as well as in the next tables, column (1) shows the regression of the survival on the covariates. Column (2) shows the regression on the covariates as well as industry dummies, (3) shows the regression on the covariates and location dummies, and (4) shows the regression on the covariates and both the industry and location dummies.

For the ISO-500 list, we can see that the log of the number of employees is positively correlated with survival, as well as belonging to a secular or religious network (compared to an unaffiliated firm which has been left out due to collinearity).

After controlling for both the industry and the city, we have a 0.4% decrease in hazard for every 1% increase in the number of employees, and a decrease of 0.14% in hazard for every 1% increase in the productivity the year of a firm's exit.

What is counterintuitive, however, is the positive correlation of the entry profitability with an increase in hazard (0.17% increase in hazard per 1% increase in profitability). Since by most analyses, the entry productivity appears to play no effect in the ISO-500 and is positively correlated with survival in the ISO0-1000 list, there does not seem to be a good explanation for this.

For the ISO-1000 list, the Table 18 displays the hazard ratios. Largely, effects are similar in the larger list, as it is in the smaller one. The 1% increase in the of employees decreases the by 0.49%. Effects of the religious and secular networks remain the same.

Table 18: Exponential survival coefficients for the ISO-1000 list

	(1)	(2)	(3)	(4)
Log the Number of Employees Before Exit Year	0.642*** (-11.61)	0.603*** (-10.51)	0.633*** (-10.89)	0.610*** (-9.53)
Secular	1.101 (0.70)	1.156 (0.92)	1.074 (0.48)	1.114 (0.63)
No Network	1.668*** (4.48)	1.611*** (3.58)	1.501** (3.23)	1.462** (2.67)
Log Entry Productivity	0.793*** (-5.86)	0.849*** (-3.36)	0.763*** (-6.12)	0.804*** (-3.94)
Log Entry Profitability	1.334*** (11.93)	1.290*** (8.83)	1.328*** (11.26)	1.292*** (8.47)
Log Exit Productivity	0.835*** (-4.92)	0.807*** (-4.66)	0.854*** (-3.61)	0.847** (-3.00)
Log Exit Profitability	0.944* (-2.12)	0.969 (-0.97)	0.960 (-1.40)	0.965 (-1.05)
Export Dummy	0.908 (-0.96)	0.871 (-1.01)	0.921 (-0.76)	0.982 (-0.13)
Significance of the Constant Industry Dummy City Dummy	*** No No	*** Yes No	*** No Yes	*** Yes Yes
N	1044	845	1044	845

Exponentiated coefficients; t statistics in parentheses
="* p<0.05 ** p<0.01 *** p<0.001"

This may be related to the higher proportion of firms that have entered religious networks after 2002, as the ISO-1000 list begins in 1997. One other interesting point is that unlike the ISO-500 list, whether a firm is an exporter or not does not appear to have a significant statistical relation to its survival. This may be because in the ISO-1000 list, 87% of the firms are exporters, which is a higher proportion to the ISO-500 list in which 76% are exporters. This, too, is counterintuitive: of the top 500 firms in Turkey a smaller proportion are exporters than the top 1000. For the rest of the covariates, the size and scale of the coefficients are similar to the ISO-500 list.

Log-Logistic Model

Next, we can look at the Log-Logistic Survival estimate in Table 18 and Table 19:

Table 19: Log-logistic survival coefficients for the ISO-500 list

	(1)	(2)	(3)
Log of the Number of Employees Before the Exit Year	0.220*** (9.34)	0.354*** (9.20)	0.340*** (8.55)
Secular	-0.102 (-0.97)	-0.103 (-0.85)	-0.0255 (-0.19)
No Network	-0.500*** (-5.43)	-0.366*** (-3.45)	-0.294* (-2.48)
Log Entry Productivity	0.0711* (2.42)	0.00647 (0.17)	-0.00270 (-0.07)
Log Entry Profitability	-0.243*** (-23.17)	-0.219*** (-17.71)	-0.216*** (-17.64)
Log Exit Productivity	0.109*** (3.92)	0.105** (2.94)	0.0884* (2.49)
Log Exit Profitability	0.0732*** (4.23)	0.0270 (1.16)	0.0203 (0.87)
Export Dummy	0.413*** (7.03)	0.389*** (3.61)	0.265* (2.37)
Constant	-1.541*** (-8.88)	-2.592*** (-9.03)	-2.915*** (-9.75)
Gamma	-0.797*** (-30.43)	-0.841*** (-23.70)	-0.923*** (-25.92)
Industry Dummy City Dummy	No No	Yes No	Yes Yes
N	1140	658	658
t statistics in parentheses ="* p<0.05	** p<0.01	*** p<0.001"	

The interpretation of the coefficients in the Log-Logistic model is slightly different from the other ones. In our results, for all the covariates except for the entry profitability, the hazard function decreases with duration, meaning that as a firm's time in the list increases, the chances of it dropping out decreases. The following are useful to note:

a) The comparative hazard ratio between belonging to a religious network as opposed to being non-networked loses much of its statistical significance, once location and industry dummies have been accounted for,
b) 'Entry productivity' loses all statistical significance once the above dummies have been accounted for.
c) 'Exit profitability' loses statistical significance once the industry dummies have been accounted for.

The coefficients are all between 0 and 1, implying that the hazard function decreases with duration with monotonicity (i.e. there is no initial period when the hazards increase followed by a decrease).

This also agrees with our other survival functions based on exponential and Weibull fits. One point to be made is that since belonging to a religious network loses much of its statistical significance once location and industry dummies have been accounted for, religious network membership should be highly correlated with certain industries and locations, firms belonging to which have high survival rates.

The ISO-1000 results are similar to the ISO-500 results and are displayed in Table 19. It resembles the exponential and Weibull functions, although the results appear to be more statistically significant for religious firms. Otherwise, the behavior of the hazard functions remains the same. Being an exporter is completely insignificant with this distribution as well.

Table 20: Log-logistic survival coefficients for the ISO-1000 list

	(1)	(2)	(3)	(4)
Log of the Number of Employees Before the Exit Year	0.380*** (13.38)	0.423*** (12.79)	0.387*** (13.24)	0.405*** (12.09)
Secular	-0.134 (-1.40)	-0.153 (-1.47)	-0.121 (-1.23)	-0.129 (-1.21)
No Network	-0.454*** (-5.83)	-0.422*** (-5.02)	-0.345*** (-4.29)	-0.308*** (-3.58)
Log Entry Productivity	0.209*** (5.65)	0.160*** (4.08)	0.225*** (6.27)	0.178*** (4.70)
Log Entry Profitability	-0.294*** (-15.73)	-0.257*** (-12.95)	-0.276*** (-15.43)	-0.251*** (-13.18)
Log Exit Productivity	0.113** (3.21)	0.129*** (3.38)	0.0996** (2.84)	0.101** (2.66)
Log Exit Profitability	0.0860*** (3.72)	0.0726** (2.99)	0.0650** (2.93)	0.0655** (2.83)
Export Dummy	0.115 (1.53)	0.129 (1.41)	0.0981 (1.32)	0.0369 (0.41)
Constant	-2.257*** (-11.17)	-2.917*** (-12.61)	-3.078*** (-13.13)	-3.514*** (-12.68)
Gamma	-0.782*** (-27.35)	-0.861*** (-27.31)	-0.858*** (-29.88)	-0.934*** (-29.45)
Industry Dummy City Dummy	No No	Yes No	No Yes	Yes Yes
N	1044	845	1044	845

t statistics in parentheses
="* p<0.05 ** p<0.01 *** p<0.001"

Gompertz Model

The final survival analysis was done using a Gompertz fit, and the results, which are consistent with the other fits, are displayed in Table 21 and Table 22.

Table 21: Gompertz survival coefficients for the ISO-500 list

	(1)	(2)	(3)	(4)
Log of the Number of Employees Before the Exit Year	0.728*** (-9.49)	0.589*** (-9.49)	0.723*** (-8.74)	0.594*** (-8.61)
Secular	0.967 (-0.23)	0.901 (-0.58)	1.042 (0.25)	0.994 (-0.03)
No Network	1.989*** (5.33)	1.749*** (3.54)	1.982*** (4.61)	1.935*** (3.42)
Log Entry Productivity	0.925* (-2.36)	0.990 (-0.21)	0.948 (-1.47)	1.055 (0.92)
Log Entry Profitability	1.256*** (15.81)	1.250*** (12.49)	1.268*** (15.72)	1.272*** (12.40)
Log Exit Productivity	0.840*** (-5.46)	0.810*** (-4.29)	0.855*** (-4.53)	0.826*** (-3.55)
Log Exit Profitability	0.944** (-2.63)	1.001 (0.02)	0.929** (-3.22)	1.018 (0.50)
Export Dummy	0.572*** (-6.92)	0.549*** (-3.78)	0.657*** (-4.65)	0.699* (-2.01)
Gamma	0.0531*** (6.67)	0.0856*** (8.59)	0.0825*** (9.79)	0.113*** (10.74)
N	1140	658	1140	658

Exponentiated coefficients; t statistics in parentheses
="* p<0.05 ** p<0.01 *** p<0.001"

Table 22: Gompertz survival coefficients for the ISO-1000 list

	(1)	(2)	(3)	(4)
Log of the Number of Employees Before the Exit Year	0.609*** (-12.60)	0.543*** (-12.05)	0.582*** (-12.49)	0.537*** (-11.37)
Secular	1.083 (0.58)	1.145 (0.85)	1.065 (0.42)	1.125 (0.68)
No Network	1.722*** (4.75)	1.731*** (4.07)	1.582*** (3.61)	1.607*** (3.29)
Log Entry Productivity	0.762*** (-6.96)	0.818*** (-4.20)	0.711*** (-7.69)	0.746*** (-5.18)
Log Entry Profitability	1.396*** (13.06)	1.381*** (10.48)	1.428*** (13.14)	1.410*** (10.42)
Log Exit Productivity	0.818*** (-5.56)	0.774*** (-5.61)	0.831*** (-4.27)	0.816*** (-3.64)
Log Exit Profitability	0.933** (-2.58)	0.961 (-1.20)	0.949 (-1.83)	0.950 (-1.48)
Export Dummy	0.901 (5.51)	0.846 (-1.22)	0.899 (-0.97)	0.955 (-0.32)
Gamma	0.0668*** (5.51)	0.106*** (7.39)	0.100*** (7.90)	0.127*** (8.70)
Significance of the Constant Industry Dummy City Dummy	*** No No	*** Yes No	*** No Yes	*** Yes Yes
N	1044	845	1044	845

t statistics in parentheses
="* p<0.05 ** p<0.01 *** p<0.001"

The behavior of the parametric estimations match the non-parametric Kaplan-Meier estimate better in the ISO-500 list which has data collected over a longer period than for the ISO-1000 list. Another type of model is the semi-parametric Cox-Proportional Hazard Model.

Cox Semi-Parametric Model

The behavior of the semi-parametric model agrees with the parametric model. Therefore, for all the models we can suggest that higher number of employees lowers the exit hazard, as well as belonging to a secular or religious network, with the secular network performing better than the religious network. The high exit productivity also lowers the result, but after controlling for all the dummies, being an exporter seems to play no effect on the survivor of the firm in the ISO-500 or the ISO-1000 lists (shown in Table 24):

Table 23: Cox-Proportional hazard rates for the ISO-500 list

	(1)	(2)	(3)	(4)
Log of the Number of Employees Before the Exit Year	0.757*** (-8.39)	0.618*** (-8.64)	0.753*** (-7.69)	0.622*** (-7.87)
Secular	0.999 (-0.00)	0.938 (-0.36)	1.053 (0.32)	1.003 (0.02)
No Network	1.903*** (4.98)	1.686*** (3.32)	1.858*** (4.20)	1.781** (3.02)
Log Entry Productivity	0.933* (-2.05)	0.984 (-0.32)	0.952 (-1.31)	1.036 (0.63)
Log Entry Profitability	1.227*** (14.00)	1.225*** (11.24)	1.236*** (13.85)	1.242*** (11.14)
Log Exit Productivity	0.858*** (-4.69)	0.824*** (-3.90)	0.870*** (-3.93)	0.842** (-3.19)
Log Exit Profitability	0.949* (-2.37)	0.997 (-0.09)	0.936** (-2.89)	1.009 (0.27)
Export Dummy	0.629*** (-5.72)	0.597** (-3.26)	0.712*** (-3.77)	0.744 (-1.66)
Industry Dummy City Dummy	No No	Yes No	No Yes	Yes Yes
N	1140	658	1140	658
Exponentiated coefficients; t statistics in parentheses ="* p<0.05	** p<0.01	*** p<0.001"		

Table 24: Cox-Proportional hazard rates for the ISO-1000 list

	(1)	(2)	(3)	(4)
Log of the Number of Employees Before the Exit Year	0.643*** (-11.19)	0.586*** (-10.61)	0.620*** (-11.04)	0.582*** (-9.98)
Secular	1.101 (0.70)	1.153 (0.90)	1.075 (0.48)	1.118 (0.65)
No Network	1.668*** (4.48)	1.648*** (3.74)	1.528*** (3.36)	1.518** (2.92)
Log entry Productivity	0.792*** (-5.77)	0.839*** (-3.60)	0.748*** (-6.46)	0.779*** (-4.43)
Log Entry Profitability	1.334*** (11.28)	1.315*** (8.99)	1.354*** (11.27)	1.336*** (8.98)
Log Exit Productivity	0.836*** (-4.83)	0.800*** (-4.82)	0.850*** (-3.72)	0.840** (-3.13)
Log Exit Profitability	0.945* (-2.09)	0.967 (-1.01)	0.959 (-1.45)	0.961 (-1.16)
Export	0.907 (-0.97)	0.860 (-1.10)	0.913 (-0.83)	0.972 (-0.20)
Industry Dummy City Dummy	No No	Yes No	No Yes	Yes Yes
N	1044	845	1044	845
Exponentiated coefficients; t statistics in parentheses ="* p<0.05	 ** p<0.01	 *** p<0.001"		

The ratios and the significance in the ISO-1000 list is similar to the proportional models as well as the ISO-500 list. Throughout the analysis for both the ISO-500 and ISO-1000 lists, the effect of entry profitability appeared to be negative. As said above, this may have been due to the high number of firms that were in the list for only 1 year, and their presence skewing the results. By and large, in the next section, when time varying covariates are added, we may able to discern the yearly effects of productivity, and profitability more accurately.

Figure 38 and Figure 39 overlay the various survival functions for the ISO-500 and ISO-1000 functions respectively on the semi-parametric model. As discussed in the Model section of this paper, the semi-parametric model separates the durations into discrete intervals and assumes a constant hazard in each one, and approximates an exponential hazard function, therefore it is natural for this semi-parametric estimate to match the parametric estimate results much more closely.

Figure 38: Comparison with the semi-parametric estimates for the ISO-500 list

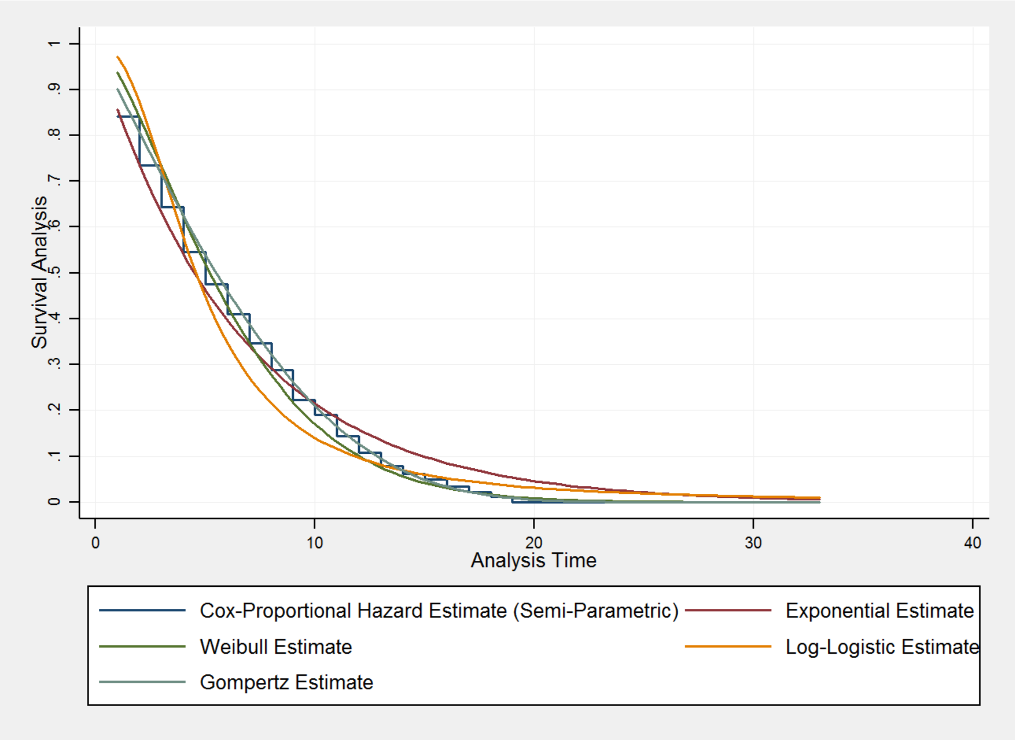
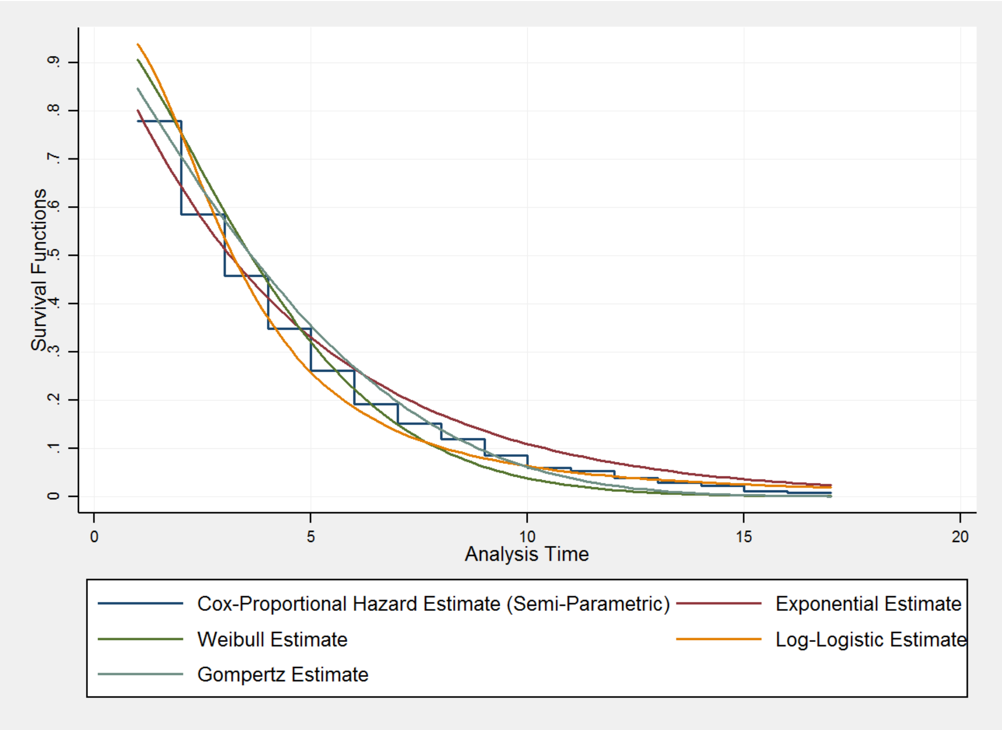


Figure 39: Comparison with the semi-parametric estimates for the ISO-1000 list



There are some controls for the validity of the Cox-Proportional Hazard Model. First is to check the proportional hazards assumption with a chi square test. In this, the null hypothesis is that the hazards are proportional in the model - that is to say, a hazard λ_i for a firm *i* can be represented as the given proportion of a baseline hazard, λ_0. Running the test for the ISO-500 and ISO-1000 data after the regression on all the covariates as well as the dummies, we have:

	chi2	**df**	**Prob>chi2**
ISO-500	67.56	90	0.9629
ISO-1000	85.81	75	0.1847

For both lists, but especially for the ISO-500 list, we have no reason to reject the null hypothesis.

ISIC Industry Codes

Table 25: ISIC codes for the various industries in the ISO lists

Code	Industry
210	Coal Mining
220	Crude Petroleum and Natural Gas Production
230	Metal Ore Mining
290	Other Mining
313	Beverage industries
314	Tobacco manufactures
321	Manufacture of textiles
322	Manufacture of wearing apparel, except footwear
323	Manufacture of leather and products of leather, leather substitutes and fur, except footwear and wearing apparel
324	Manufacture of footwear, except vulcanized or moulded rubber or plastic footwear
331	Manufacture of wood and wood and cork products, except furniture
332	Manufacture of furniture and fixtures, except primarily of metal
341	Manufacture of paper and paper products
342	Printing, publishing and allied industries
351	Manufacture of industrial chemicals
352	Manufacture of other chemical products
353	Petroleum refineries
354	Manufacture of miscellaneous products of petroleum and coal
355	Manufacture of rubber products
356	Manufacture of plastic products not elsewhere classified
361	Manufacture of pottery, china and earthenware
362	Manufacture of glass and glass products
369	Manufacture of other non-metallic mineral products
371	Iron and steel basic industries
372	Non-ferrous metal basic industries
381	Manufacture of fabricated metal products, except machinery and equipment
382	Manufacture of machinery except electrical
383	Manufacture of electrical machinery apparatus, appliances and supplies
384	Manufacture of transport equipment
385	Manufacture of professional and scientific, and measuring and controlling equipment not elsewhere classified etc.
390	Other Manufacturing Industries
410	Electricity, Gas and Steam
	Mining
	Low-Technology
	Middle Technology