Determinants of venture capitalists' exit strategies: An empirical study through survival analysis

Auteur : Gillain, Axel
Promoteur(s) : Surlemont, Bernard
Faculté : HEC-Ecole de gestion de l'ULg
Diplôme : Master en ingénieur de gestion, à finalité spécialisée en Financial Engineering
Année académique : 2015-2016
URI/URL : http://hdl.handle.net/2268.2/1845

Determinants of venture capitalists' exit strategies: An empirical study through survival analysis

Jury :
Promoter :
Bernard SURLEMONT
Reader(s) :
Ashwin ITTOO
Olivier de WASSEIGE

Dissertation by
Axel GILLAIN
For a Master Degree in Financial Engineering
Academic year 2015/2016

Acknowledgments

I would like to express my gratitude to my promoter Bernard Surlemont for the guidance as well as the useful comments, remarks and engagement through the learning process of this master thesis.

I would also like to thank my readers Ashwin Ittoo and Olivier de Wasseige for their availability.

Last but not the least, I would like to thank my friends and family for their support and precious help.

Executive Summary

In this dissertation I study the determinants of venture capitalists' exit strategies, and more specifically, the interactions between exit type and timing. Indeed, in addition to knowing how they plan to exit, venture capitalists are also interested in knowing when they will be able to do so. Examining the exit strategies of venture capitalists thus requires to take those two dimensions into account.

Through the use of survival analysis methods, I analyze a sample constituted of more than 19.000 financing rounds in 11.500 unique firms. Set in the framework of competing risks models, this rigorous statistical analysis gives some interesting insight on the relationships between a series of variables (such as the stage at which the round takes place, the syndication of the deal, the industry of the firm, and so on) and the time needed for an exit to occur.

Moreover, when considering the type of investor I make the distinction between business angels and venture capitalists. This is therefore the first time that the impact of business angels on the exit strategies of venture capitalists is studied using survival analysis methods.

The results show that the presence of business angels allows firms to exit through acquisition both faster and more often. However, business angels do not seem to have a meaningful impact on the likelihood of liquidation.

Furthermore, it can also be concluded from the analysis of the results that the benefits from deal syndication are real. Indeed, when at least two venture capitalists are present, investments appear to exit through acquisition substantially more often and up to 26% faster. Liquidations are also significantly less likely to occur when the number of venture capitalists involved increases.

Table of Contents

Table of Figures

I. Introduction

Venture capital is a well-known source of funding for businesses with high potential for growth but also bearing an amount of risk significant enough to scare off banks. Even though venture capital funding is a relatively new industry that only started to boom a few decades ago, it has reached today unprecedented levels, with almost $60 billion invested across the United States in 2015[1].

And as the sector has grown in size, it has become a very well documented topic. Today, a search on Google Scholar (a web search engine that indexes scholarly literature) with the terms "venture capital" yields almost 1.5 million results. In addition to the profuse number of scientific papers on the subject there also exist a vast number of guides and handbooks for the finance-seeking entrepreneur as well as for the wannabe investor (e.g. Bygrave, Hay and Peeters, 1999; Bloomfield, 2008).

The process of venture capital financing, for both the investee and the investor, is therefore well known and many papers have already examined the topic. An aspect of particular importance for the venture capitalist in the midst of his investment strategy is the evaluation of possible exit routes. Indeed, it is widely recognized that the decision to "enter" an entrepreneurial venture is based on the exit possibilities. If there is no chance that the investment will become liquid after some time, in other words if the venture presents limited exit possibilities, the venture capitalist will usually not invest.

Furthermore, not only do the investors need to evaluate how they will be able to exit their investment, they also need to evaluate when they will be able to do so. The two dimensions of the exit strategy – the duration and the form – are interconnected and should therefore be considered simultaneously.

Even though there already exists plenty of literature on both the type of exit – and particularly on the Initial Public Offering ("IPO") – and the timing, relatively little has been written on the interactions between the two.

Regarding trade sale, i.e. when a company is acquired by another, one reason why only little empirical research has been done on the topic of trade sales is the particular

[1] According to the "Q4 2015/full year 2015 summary" MoneyTree Report by PricewaterhouseCoopers and the National Venture Capital Association, based on data provided by Thomson Reuters.

difficulty of collecting data on this type of exit. Unlike with IPOs, a firm[2] is under no obligation to publicly disclose information when undergoing a trade sale. The same lack of information affects the write-offs (liquidations). Here, it is even worse since liquidations are the type of exit no one will want to advertise, which may explain why few papers have included it in their considerations of possible exit routes (Schwienbacher, 2009; Bloomfield, 2008).

Share buybacks and secondary sales are other possible exit routes, in which venture capitalists sell their shares back to the management or to other institutional investors. These types of exit are typically associated with "partial exits", meaning that not all shares are sold at the same time (Cumming and MacIntosh,2003a). One common cause of partial exits occurring more often in the case of share buybacks or secondary sales is the recording of poor results by the company, resulting in low returns for the investors. Therefore, in order to certify that such companies are still valuable, investors will choose to remain partly involved by keeping only some limited financial commitment (Schwienbacher, 2009).

This study thus builds on previous research that used survival analysis (e.g. Gompers, 1995; Giot and Schwienbacher, 2005; Félix et. al, 2012) in order to model the exit strategies of venture capitalists by adding a new variable: the presence of business angels amongst the investors.

To my knowledge, this is the first time that the impact of business angels on the exit strategies of high-tech startups has been studied through survival analysis. Most authors who have included characteristics related to venture capitalists ("VCs") in their empirical studies have focused on the impact of variables such as age, size, reputation, network, location, contracting or monitoring policies, syndication, affiliation to financial institutions, and so on, but have never made the distinction between business angels and venture capitalists.

A. An innovative database

The distinction between the two types of investors is made possible thanks to CrunchBase, a relatively new database operated since 2007 by TechCrunch, one of the most highly regarded web publishers of news on the technology industry. The CrunchBase database contains around half a million data points listing companies, people, funds, funding rounds, events, and details on each element and their relationship between each other. Anyone, after registering, can make submissions to the database; however, any change is subject to review

[2] In the United State

by a moderator (and usually supported by a news article) before being validated. Data is constantly reviewed by editors to ensure it is up to date. Overall, CrunchBase mostly profiles, but is not limited to, companies in the Internet industry.

So far the CrunchBase database has not been a very popular source of data for empirical studies. However, Block and Sandner (2009) used CrunchBase data on Internet start-ups for their paper and found a significant correlation between this database and the data provided by the National Venture Capital Association.

Among the data available on the website are details of the funding rounds undertaken by companies, such as the date of the round, the investors involved (if publicly available), the amount, etc. Therefore, in addition to the type of investor, CrunchBase allows the study of a wide range of other variables: the stage at which the financing round occurs, the amount invested, the sector, the number of investors, the number of previous rounds, or the duration of each round. Interestingly, all of this information is known to the investor at the time of the funding round. This also allows the analysis of the exit conditions that evolve as firms progress in their financing rounds and to contrast these results with the conditions at the time of the initial investment.

Furthermore, since the database has a focus on the Internet industry, it gives the possibility to examine whether these variables show the same behavior as they did in other research that studied a broader range of industries. Differences from the literature might indicate specificities of the Internet sector with regard to certain characteristics.

Unfortunately, CrunchBase does not give any information about share buybacks and secondary sales even though these are possible exit routes. For the rest of this paper, share buybacks and secondary sales will not be discussed in much more detail. The same goes for partial exits, since there are no mentions of them on the website.

Although all the up-to-date information is freely accessible through CrunchBase's website, access to the most recent database is restricted to premium users. Fortunately, CrunchBase gives the possibility to use a free snapshot of its data collected before December 2013. Using this information I was able to extract more than 19.000 investment rounds in 11.500 unique firms that will constitute the dataset on which I will be able to model the exit strategies of venture capitalists.

B. Objectives

To get the most out of the data, the most suitable statistical analysis tool should be used. In this paper the objective of the empirical study is two-fold; to determine which variable has an impact that is statistically significant on the exit strategy, and to measure the impact of each statistically significant variable on the variables of interest. The variables of interest (dependent variables) when studying VCs' exit strategies are:

- The type of exit: IPO, acquisition or liquidation,
- The duration of the funding rounds

Linear regressions are not suitable since they cannot have dependent variables that are categorical (the exit type is either IPO or acquisition or liquidation, each event being mutually exclusive). Logistic regressions, on the other hand, use binary dependent variables, but would not give any information about the time factor. It would be possible to implement a model using several logistic regressions – a binomial model using a logistic regression for each type of exit – but this would not allow the analysis of the interactions between each type of exit. Furthermore, logistic regressions do not integrate the information provided by investment rounds that have not been exited at the time of the study. The best type of analysis is thus the survival analysis, since it allows categorical variables through multiple risks models, and is capable of incorporating the partial information provided by an investment that has not been exited.

Survival analysis analyzes the expected duration of time until an event happens (time to event). The term "survival analysis" comes from its initial area of studies where the event of interest was death: it was – and still is – commonly used for clinical trials. Nowadays its scope has become much broader and survival analysis is used in a wide range of different fields and applications (Singh and Mukhopadhyay, 2011). In this case, since there are multiple mutually exclusive events, I set the framework of the survival analysis using competing risks models. The advantage of this analysis is that it makes it possible to estimate for each type of exit the significance and the impact of each variable on the survival time.

Throughout this paper the following research questions are analyzed in detail:

1) What are the factors related to the company and the funding round that impact the outcome of an investment? I focus on characteristics of the venture (e.g. the financing stage it has reached or the industry type) and of the funding round (e.g. number of participants, investor type or amount raised) and not on the

characteristics of the investors (e.g. reputation or network). However, the distinction between angel investors and VC is made.

2) What is the impact of the previously identified factors? Does any factor increase or, on the contrary, reduce the time to exit? What is the significance for the time to exit of the increase or decrease in one of the studied factors?
3) Does the presence of venture capitalists or business angels amongst the investors of the firm have an effect on the exit route and timing of the firm? Which of these types of investor has the strongest impact?

Regarding question (3) I formulate the following hypothesis:

a) The presence of business angels or venture capitalists influences the exit route and timing.
b) Their presence reduces the time until a favorable exit occurs (IPO or acquisition) and increases the time until an unfavorable exit (liquidation).
c) The likelihood of a successful exit is increased when business angels are involved.

C. Limitations

Although one of the objectives of this paper is to identify significant factors explaining the exit route and timing, the aim is not to identify every possible factor. There are two reasons for this.

First I am limited to those variables that are available through the database. Unfortunately there are no straightforward methods for collecting each and every piece of information that may or may not be related to the exit strategy. For example, some studies have shown that more experienced and more reputable venture capitalists have a better ability to time their exit when markets are optimal (among others: Lerner, 1994a; Gompers, 1996). And while it is possible to know exactly which business angel or which venture capitalist is involved in a financing round through CrunchBase, there are no direct variables that capture the experience or the reputation of an investor. It might be possible to proxy this with, for example, the number of investments they have realized (age would be another possibility but unfortunately, the database does not give this information). But then the database would be misleading since the only investments that would be counted would be those that had been added to CrunchBase, instead of the actual number of investments. After all, some investors might be more popular than others among the users of the website, and thus would have their investment scrutinized in a more thorough manner, leading to a greater presence in the database than other less popular investors. Therefore, the estimate of investors' experience would be biased by their popularity among CrunchBase's users.

Second, regardless of the database, it is a vain enterprise to try to identify and measure all the characteristics of individuals, companies and their environments, since no set of measured variables can possibly capture all the variation among them. For example, two ventures may be similar in all respects but just having two different entrepreneurs may dramatically change the outcome that they will experience. In an ordinary linear regression model, this unobserved heterogeneity is represented by a random disturbance term. When considering a typical linear regression

$$y_i = \boldsymbol{\beta}_i x_i + e_i \tag{1.1}$$

where e_i represents all unmeasured sources of variation in y_i. In such a model, it is typical to assume that e_i has a normal distribution with a mean and variance that is constant over i, and that the e's are independent across observations (e is an independent and identically distributed (iid) random variable). In survival analysis, the log-normal accelerated failure time (AFT) model has exactly these assumptions. Other AFT models give the possibility to specify distributions for e besides the normal distribution but retain the assumptions of constant mean and variance (Allison, 2010), as well as independence across observations (such alternative models are considered in Chapter 5). The impact of unobserved heterogeneity will be considered in more details later on.

This rest of this paper unfolds as follows. In Chapter 2 the theoretical foundation on which this paper builds is presented through a thorough literature review. Then, in chapter 3, I detail the variables that will be investigated and their expected impact on the duration until exit. Chapter 4 gives more information about the data sample on which the empirical study is based and its validity. In Chapter 5 I define the methodology used and more precisely the bare bones of survival analysis and multiple risks models. Chapter 6 gives detailed descriptive statistics on the sample and a discussion on the estimation results. Finally, chapter 7 concludes the paper.

II. Literature review

A. Initial Public Offering

Among the possible exit routes, the IPO is undoubtedly the one regarded as the most successful[3] and the one associated with the highest returns, for the investor as well as the entrepreneur[4]. This may explain why much of the previous literature has focused on it. In addition, the accessibility of the information on IPOs simplifies data gathering for empirical studies.

Numerous studies have focused on the timing dimension of the IPO exit. For example, one way for venture capitalists to time their exit is by using stage financing. Ruhnka and Young (1987) define a "venture capital model" in which venture capitalists distinguish a number of stages based on the characteristics of ventures in each stage, key developmental goals or benchmarks typically accomplished in each stage, and the major risks involved.

In each stage, the entrepreneur gets the required funding to proceed to the next development stage, but venture capitalists refrain from giving more money than actually needed. Gompers (1995) argues that venture capitalists use stage financing to regularly monitor the firm and thus keep the option of discontinuing funding projects with little probability of going public. Additionally, Gompers (1995) finds that there is a positive relationship between the duration of each financing round and the tangibility of assets, and a negative relationship between the market-to-book ratio and the R&D level (variables that are related to the intensity of asymmetric information). Bergemann and Hege's (1998) model shows that time-varying contracts, such as finance staging, are optimal because they provide knowledge to impede the asymmetric information problem.

Cumming and MacIntosh (2001) establish a theoretical model using survival analysis to investigate the optimal duration of the venture capital investment. They use the model to predict the theoretical effect of a few factors (the stage of the venture at the time of the first investment, the size of the venture capital industry at the time of the investment, whether exit was planned or not and whether the exit is on response to an unsolicited offer) on the investment duration. Their study shows that an increase in the availability of venture capital has a negative impact on the duration. Unlike Gompers (1995) who studied the duration of

[3] See, among others, Lerner (1994), Black and Gilson (1998), Flemming (2002), Bascha and Walz (2001), Cumming and MacIntosh (2001, 2003a,b), and Schwienbacher (2002)

[4] See Bygrave and Timmons (1992), Gompers and Lerner (1999b), and Cochrane (2001)

each financing round, Cumming and MacIntosh (2001) investigate the total span of the venture capital investment. However, neither Gompers (2005) nor Cumming and MacIntosh (2001) set their analysis in the context of competing risks.

Lerner (1994a) explores the relationship between VCs' reputation and the timing of their IPOs and finds that reputable VCs are able to a greater extent to time their exit when stock markets are peaking and to use private financing to grow the firms privately when markets are down. Gompers (1996) shows that there exists a relationship between venture capitalists trying to establish their reputation and the timing of their IPOs. Younger, less experienced and less reputable venture capital firms have a shorter time to exit because of their need to signal value in order to attract investors for follow-on funds. As a result they may not go public when the markets are optimal.

Black and Gilson (1998; 1999) argue that venture capital thrives especially when venture capitalists can exit from successful portfolio companies through initial public offerings (IPOs), which in turn requires an active stock market. Michelacci and Suarez (2002) suggest that the public markets develop innovation, business creation and growth by giving the possibility for "informed capital" to recycle their investment. Furthermore, the sooner new ventures go public, the faster this informed capital can be redirected toward new companies.

Others have studied the relationship between VCs' experience, reputation or network and exit strategy, for example Nahata (2008) finds that companies backed by more reputable VCs are more likely to exit successfully, access public markets faster, and have higher asset productivity at IPOs. Hochberg, Ljungqvist and Lu (2007) find that better-networked VC firms experience significantly better fund performance, as measured by the proportion of investments that are successfully exited.

Das, Jagannathan, and Sarin (2003) estimate the probability of investment rounds ending by acquisition or IPO and found, in their sample, a cumulative probability of exit by IPO of 20-25%, more of less consistently over the different stages of financing. Their estimated probability for exits by trade sale is between 10 and 20%.

B. Other exits

However, an equally important exit option for investors to accomplish the same objectives is agreeing to acquisition by another firm.

According to Poulsen and Stegemoller (2008) IPO firms are larger and more profitable firms; VC-backed firms are more likely to go public than to be acquired. For Cumming (2008), financial contracts that give the venture capitalist greater control over the governance of the firm increase the likelihood of the firm being acquired rather than going public. Bayar (2006) and Poulsen and Stegemoller (2008) also find that firms characterized by higher pre-exit sales growth are more likely to go public rather than be acquired.

Brau, Francis, and Kohers (2003) show that an "IPO valuation premium" exists by comparing the valuation multiples of IPOs and acquisitions, and find that sellers in acquisitions receive payoffs equaling just 78% of those in IPOs. Their results indicate that the concentration of the industry, high-tech status, the liquidity of the IPO market relative to the acquisition market, the percentage of insider ownership and firm size are all positively related to the probability that the firm will conduct an IPO. On the other hand, private companies in high book-to-market industries, firms in the financial service sector and other highly leveraged industries and deals involving greater liquidity for selling insiders show a greater likelihood of being acquired.

As opposed to IPOs, trade sales can be seen as a more universal exit route open to many companies and particularly to the less successful ones. Das, Jagannathan, and Sarin (2003) find that companies in the later stages of development are more likely to be acquired. They argue that this may be because many firms that cannot reach the IPO stage conclude a trade sale instead. They also find a relationship between the time to exit and the stage of financing. Favorable exits in their study occurred within a three year period after the financing round for over two-thirds of late-stage companies. This proportion dropped to only one-third for early-stage companies (still within three years of financing). Besides, they establish a relationship between the gains of venture capital investments and variables such as the amount invested, the industry, the valuation at the time of the funding round, and the market sentiment. They also conclude that high-tech investments have a higher probability of achieving a favorable exit with success.

Schwienbacher (2008a) also supports the idea that IPO is an exit that may be limited to the most promising ventures whereas acquisitions appear to be a more general exit route, i.e., for both more and less promising ventures. The choice of exit route for venture-backed companies is influenced by the number of financing rounds, the investment duration and the reporting requirements of the investee.

Since venture capitalists tend to invest in high-tech industries (Brau et al., 2003) such as software, electronics and hardware, Internet and telecommunication, and biotechnology, the funding through venture capital for high-tech startups has also been quite well researched. Venture capitalists are an important source of capital for innovative high-tech start-up firms (Gompers and Lerner, 1999a). Around 90% of all venture capital funding in North America is in technology firms (Cumming, 2007). Authors such as Murray and Lott (1995) and Lockett et al. (2002) have shown that VCs may be hesitant to invest in early-stage high-tech startups. The lack of funding for early-stage high-tech companies has typically been referred to as "the equity gap" (Murray, 1998). Moreover, Schwienbacher (2008b) analyzes how venture capitalists's exit preferences (IPO or trade sale) influence the innovation strategy of startups.

Other exit routes are also possible. For example, venture capitalists can liquidate their investments through secondary sale (when only the venture capitalist sells his shares), buyback (when the entrepreneur repurchases the shares), or write-off (liquidation). A partial exit for each route is also possible (Cumming and MacIntosh, 2003a).

Though numerous aspects of the exit strategy have been documented, relatively few researchers consider all the possible exit routes simultaneously, and none of the previous papers take into account the possibility of a liquidation in their study.

Schwienbacher (2002) investigates the exit types in the US and in European countries taking into account all possible exit routes. He examines the impact that VC-related variables and monitoring policies have on the choice of an exit route. His analysis shows that significant differences exist between the US and European venture capital markets. His results also show that the monitoring policies have an impact on the probability of the exit by IPO.

Cumming (2008), Schwienbacher (2002), for example examine the whole range of exit possibilities, but these studies only analyze the type of exit; they do not consider the timing of the exit or the interconnections between timing and type of exit.

Only a handful of papers have studied the interrelations between the two dimensions of the exit strategy, in other words how the various exit routes interact with each other over time. For example, Giot and Schwienbacher (2005) study the period of the investment and consider each kind exit of route simultaneously. Using survival analysis they model the time to exit (like Gompers, 1995, and Cumming and MacIntosh, 2001). Moreover, since they analyze various forms of exit, they use a competing risk model. According to their analysis, the different forms of exit may have different conditional exit rates (hazard functions).

Another conclusion they reach is that the conditional exit rate is not monotonous (not always either increasing or decreasing), and the hazard rate varies relative to the exit type. Their results also show that the industry of the investee has an impact on the timing of the exit (e.g. biotech and Internet firms have the shortest investment duration before an IPO, and biotech firms are those that take the most time to liquidate).

Félix, Pires and Gulamhussen (2012) investigate the total investment duration and the impact on exit strategy of the characteristics of venture capitalists, of their investments and contracting policies. They highlight the importance of variables related to venture capitalists' financial contracts and monitoring (reporting requirements, venture capitalist presence on the board of directors, syndication percentage) and find that the association of the venture capitalist with a financial institution leads to a shorter investment duration for all types of exit. Their results also indicate that hazard functions are non-monotonic for all exit forms.

Ozmel, Robinson and Stuart (2012) use survival analysis to examine the relationships between alternative funding sources in the private capital market and the startups exit strategy in the biotech industry. However, they do not take into account the possibility of a write-off occurring before the firm's exit from the private capital market. Their results show that strategic alliances (inter-firm commercialization agreements) and venture capital funding increase the hazard of going public as well the hazard of being acquired for startups. Moreover, they show that biotechnology firms that have VC investments from better networked VCs with more central positions in VC syndicate networks exit significantly faster through IPO.

C. Business angels

So far, a number of academic papers dealing with venture capitalists have been reviewed, but as previously mentioned, this study also highlights the potential differences between the investment process of business angels and venture capitalists, more specifically during the post-investment stage. It will be useful to begin by stating the differences between the two.

The literature on business angels recognizes that investors of this type tend to appear earlier in the life of the startup, that the funds they provide are in lower quantities and usually represent the first form of equity funding for the investee, and that they also contribute advice, and networking opportunities as well as providing more hands-on assistance to the entrepreneur. This may be due to the fact that, while venture capitalists invest other people's money, business angels invest their own.

For these reasons, business angels fill a gap left by the venture capitalists in the financing of very early-stage startups. For a very young startup, the funding of its very first product may be incompatible with the traditional venture capital investment model in several respects. For example, venture capital funding is often preceded by a lengthy diligence process, the amounts needed by the new ventures are far below what venture funds seek to invest, and venture capitalists typically require board representation, which can be problematic when the firm is at a stage when it does not even have a board. As a result, venture capital requires a high level of commitment from the entrepreneur to pursue a project that may be extremely experimental.

Angel investors are thus an interesting option for entrepreneurs in order to bridge the gap between building the initial product and building the company. Furthermore, angels typically expect to take a role that enables them to contribute not only strategically but also operationally (Paul et al., 2007).

Most of the literature investigates the role and impact of venture capitalists on entrepreneurial firms. By contrast to the literature on venture capitalists, Denis (2004) reports that "comparatively little work has been done on angel investors".

Some studies, however, have investigated the impact of business angels on the successive financing rounds of a startup. For example, Madill et al. (2005) found that the majority the firms they questioned received venture capital after receiving business angels' money. They suggest that "angels help the new venture to become more ready for future investment by, among other contributions, being closely involved with the firms in which they invest" (Madill et al., 2005), and conclude that angels significantly increase the appeal of businesses to venture capitalists, by demonstrating a track record of performance to such potential investors.

Similarly, Kerr and Schoar (2010) find that angel-backed firms are significantly more likely to survive for at least four years and to raise additional financing outside the initial angel group. Such firms are also more likely to show improved performance and growth (measured by growth in web traffic and website rankings). The improvement gains typically range between 30 and 50%.

Bloomfield (2008), and Wong et al. (2009) argue that angel financing comes earlier than VC financing and the funds invested are smaller. According to Wong et al. (2009), angel investors start providing financing 10.5 months after the creation of the business on average. Denis (2004) suggests that the information asymmetry and the moral hazard problem make it difficult for new ventures to achieve external capital.

Venture capitalists, however, have an advantage over angel investors in overcoming such problems. Since they invest later, they benefit from information about the viability of the business and the use of previously obtained funds (possibly from angel investors). Typically, this information is unavailable during the initial start-up phase. These problems might even be so important for some firms that venture capitalists may not even consider them as an investment opportunity. Therefore, entrepreneurs may have to explore other financing options for the initial startup stage; informal financing is one option that may be preferred by entrepreneurs (Vos et al., 2007).

The most frequent investors in early-stage ventures are business angels, who, according to Sohl (2011), invest in twenty times more ventures than venture capitalists. Furthermore, not only do business angels invest more often than venture capitalists, and more money (Sohl, 2011), but they also invest at earlier stages in the venture creation process. Venture capitalists prefer investing later in the finance cycle in order to benefit from shorter time to exit and lower perceived risk (Sapienza, Manigart, & Vermeir, 1996).

Evidence about the exit strategies of angels is limited but suggests that a majority have no preference between a trade sale, initial public offering or another type of exit (Paul et al., 2003). But there is also the possibility that the venture will fail or only be a moderate success in which case it may be difficult for the angel to exit, what Ruhnka et al. (1992) call a “living dead” investment.

This section covered the main aspects that were studied by the academic literature; the next section will also detail some previous findings which are specifically related to the variables that will be incorporated in the model.

III. Variables

This section details how the companies in the dataset and the financing they obtained from venture capitalists are described through the variables (called covariates) that will be used throughout the analysis.

With the exception of the variables of interest – the duration and exit type – all other variables are known to the investor at the time of the funding and are either related to the firm or to the ongoing financing round, apart from those describing the state of the IPO market.

More details about the academic literature on these variables are also provided, as well as intuitive expectations about their behavior in the specific framework of the analysis.

A. Firm-related variables

Firm-related variables are used to describe the company. They are fixed at the time of the financing round, but they may change throughout the financing cycle of the firm to reflect the evolution of the company.

1. The industry type

As mentioned in the introduction, CrunchBase more specifically references web-related companies, which is not surprising since the data present on the website are edited by the users themselves. To capitalize on this, I will also target this industry in the empirical study by discarding all financing rounds that do not involve a high-tech company in the Internet industry (e.g. biotechnology, pharmaceuticals, retail, healthcare, insurance, semiconductor, industrial, etc.).

Web-related companies are defined as firms with a business model that is fundamentally dependent on the Internet. Even though the company's business model is based on the Internet, the services and products it offers can be categorized.

The type of industry of the firm is thus described through a set of 12 dummy variables: ADVERTISING (the advertising industry), ECOMMERCE (the e-commerce industry, e.g. Amazon), EDUCATION (the education industry, e.g. Wikipedia), ENTERPRISE (the industry of services and products to enterprises, e.g. SAP, Zimbra, Bloomberg), GAMES_VIDEO (the video game industry), HARDWARE (the hardware industry, e.g. Intel, Samsung), MOBILE (the mobile phone and related products industry, e.g. Android, AT&T), NETWORK_HOSTING (net cloud computing, network and other hosting services, e.g.

Comcast, Dropbox,), SERVICE (the personal services industry), SOCIAL (the social services industry, e.g. Facebook, BuzzFeed), SOFTWARE (the software industry, e.g. Skype, Microsoft), WEB (the web industry, e.g. Youtube, Ebay, Google). These variables are equal to 1 (0) if the firm belongs (does not belong) to the specified industry.

2. The financing stage

While it is generally recognized that he funding needs of a company follow its development, the definition of the funding and development stages of a startup vary significantly depending on the source of information used. This section briefly defines each development stage and its financing as well as the construction of the variables.

In their paper, Ruhnka and Young (1987) define five stages of development for startups based on the typical characteristics of ventures in that stage, the developmental objectives or benchmarks usually concluded in that stage, and the major associated risks. Those stages are “seed”, “start-up”, “second stage”, “third stage”, and “exit stage”. Although CrunchBase defines the funding process of startups using other terms, the global concept remains the same, as depicted in Fig 1.

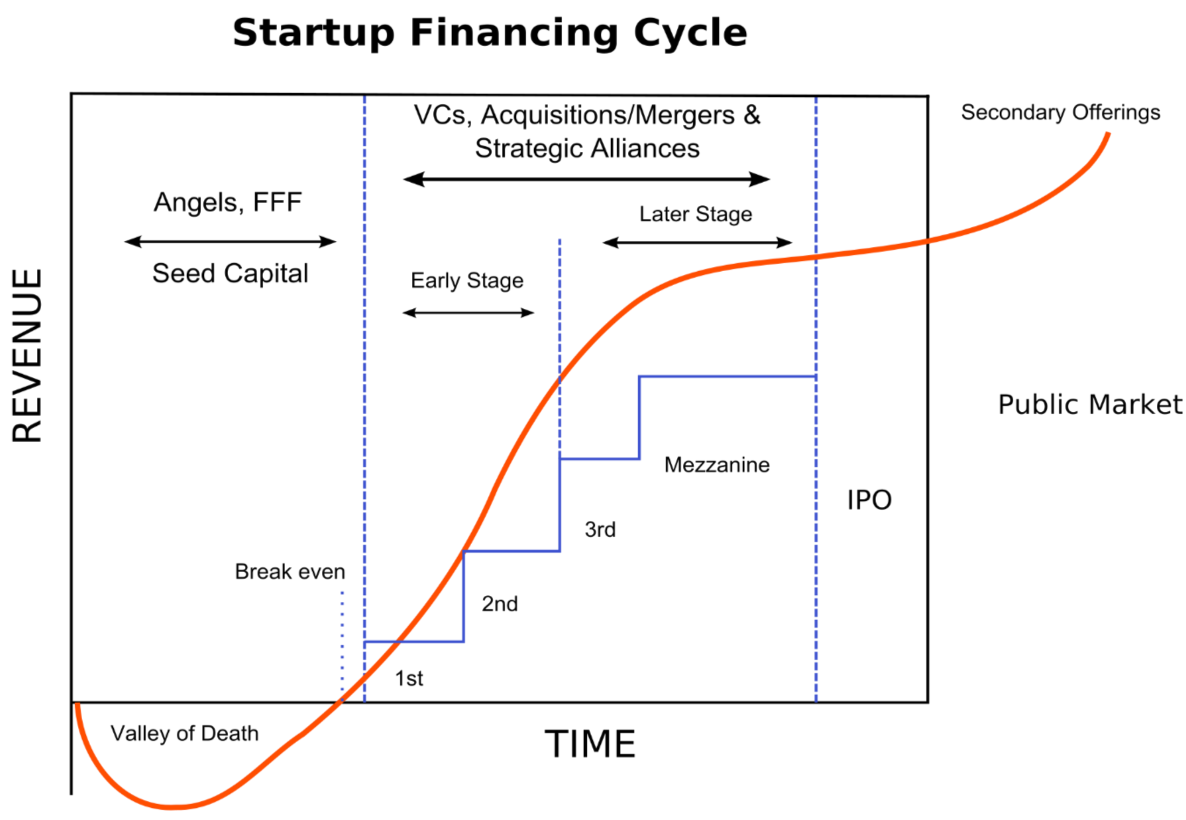


Fig 1: the startup financing cycle

The earliest possible funding for a startup is *seed capital*. This is a very early investment that can be used to pay for preliminary operations such as market research, and product development, meant to support the business until it can generate cash of its own, or until it is ready for further investments. The investor typically investigates the technical and economic feasibility of the idea. The options for raising seed capital include: the three Fs (friends, family and fools), crowd funding, angel funding, early-stage venture capitalists, or government programs (soft loans, grants, etc.).

After raising seed capital, and if the startup continues its development, it will enter the so called *early stage*, and have the possibility to raise more capital through a “series A” round. A series A round is the name usually given to a company's first significant round of venture capital financing. It refers to financing that is intended to finance the company for six months to two years, and the amounts committed vary widely depending on the firm (companies in industries that are associated with higher fixed costs, such as semiconductors, real estate or pharmaceuticals, usually raise more money than others). At this stage the firm generally develops its products and infrastructure, hires its initial employees, and undertakes early-stage business operations. This financing stage typically involves professional investors such as early-stage venture capital firms, and angel investors.

Once the series A round has been raised the firm can continue raising money through several successive financing rounds, i.e. a series B round, a series C round, a series D round, etc. However, the rounds are not necessarily named sequentially A, B, C, etc., as they happen. The letter name of the round refers to the stage that the company is in, whether it is budding or growing. A company could, for example, raise two successive series A rounds, or jump from series A to series C.

When leaving the early stage, the company is usually considered to be entering the *expansion stage*, or growth stage, and will seek to develop its operations, for example by scaling its business model, its user or customer base, or by making acquisitions. This stage is often associated with a series B round. The amount raised can range from a few millions to tens of millions, and will be invested by venture capital firms or firms specialized in later-stage investment deals.

Finally, in the *later stage,* when companies reach some degree of maturity, they usually accelerate what they have been doing in the previous stage, i.e. continue to grow fast, internationalize their operations, or acquire other companies. The financing rounds at this stage, series C onwards, tend to range from tens to hundreds of millions. A difference between the financing rounds at this stage and the previous, besides the amount being invested, is that at this point private equity firms and investment banks tend to be the lead investors, with the participation of large venture capital firms. From this stage on the outcome tends to be an IPO or aquisition by a much bigger company.

Unfortunately, the series denominations can be ambiguous. For example, a funding round can be categorized as series A because it is the first time that the company has raised venture capital, whereas it should be regarded as series B because of the amount involved and the use that will be made of the money. Furthermore, there is no obligation, for either the company or the investors, to define the funding round. For this reason, many funding rounds present in the CrunchBase database do not specify the series, just the source of the money (i.e. "venture", "angel", "private equity", etc.).

Four dummy variables are thus defined to characterize the development of the company (set to 1 (0) if the financing stage matches (does not match) the description of the variable):

- SEED: for financing rounds defined as "seed" or "angel" on CrunchBase as well as for undefined rounds with an amount lower than $1.5 million.
- EARLY: for series A rounds and undefined rounds with investments from $1.5 to $3 million.
- EXPANSION: for series B and C rounds and undefined rounds with investments from $3 to $15 million.
- LATER: for any other financing round.

It has already been mentioned that several authors have found that the staging of finance is optimal for venture capitalists, especially because it gives them the possibility to stop funding those ventures that do not reach their objectives (Gompers, 1995, Bergemann and Hege, 1998). Entrepreneurs thus have incentives to perform and show results to investors if they hope to raise more capital in the future.

It is clear that the earlier the stage, the less advanced the company and the further it is to an exit. It should thus not be surprising to see the results of the empirical study showing a positive impact of the stages of development on the time to exit. Besides, because early-stage companies are further away from their potential exit, they carry more risks for the investors.

The survival rate of the early-stage company should then be lower than that of those at a later stage, and the risk of failure should decrease as the company advances.

3. The number of previous rounds

In order to have additional information on the funding of the company, a variable that gives the ordinal number of the financing round is added (which is not specified by the financing stage variable): ROUNDNB.

It seems clear that there should be a positive relationship between the number of rounds and the survival rate, as failing businesses do not have the opportunity to raise capital. Furthermore, companies with more financing rounds should be those that are the most prone to exit successfully.

As discussed during the literature review, stage financing gives the possibility to investor to cease funding companies that have little chance of success. Conversely, companies that consistently attract capital are thus seen as having a higher chance of attaining a successful exit. This builds directly on the motives of venture capital investors. An investor provides capital to a startup with the aim of later exiting the investment, in the form either of an IPO or of a sale to another firm (Gompers and Lerner, 1999a). Each funding round is therefore expected to reduce the time until an IPO or a trade sale is made, and increase the time before a liquidation.

4. Development milestones

The realization of objectives is important for any company, but for growing companies it can be particularly critical, since achieving developmental milestones signals the firm’s quality to potential investors. A company that advances through its development stages faster than another could be seen as more successful.

The achievement of milestones is not limited to development stages. It can also be related to progress in any field that may be deemed crucial for the success of the company. This could include, reaching a critical number of users, releasing a product, realizing a technological advancement, winning an award, developing a new prototype, surviving the release of a new product, signing a partnership, and so on.

A company reaching a development milestone can hope for its progress toward a favorable exit to be accelerated. A positive impact of the number of milestones achieved on the time to a successful exit would therefore be expected, and a negative impact on the time to an unfavorable exit.

MILESTONES_ROUND: gives the number of milestones achieved up to the actual funding round.

This variable, however, is to be treated with caution since, unsurprisingly, popular startups tend to attract more attention and thus see their profiles completed much more accurately than others. One should consider that there is a possibility that this variable may not be completely unbiased.

5. IPO Markets

A number of studies have shown that the level of activity of the IPO markets has a significant impact on VCs' exit strategies (see among others Black and Gilson, 1998; 1999).

When markets are optimistic about the state of the economy and confident in their expectation of strong results, i.e. when markets are bullish, investors are more prone to buy newly issued stocks. This facilitates the exit of successful startups and leads to a decrease in the investment time (Gompers, 1995). For example, the frenzy of the Internet bubble pushed a fair number of companies to go public that later appeared to be rather questionable businesses. However, for unsuccessful companies this could also mean that investors may be hastier in abandoning their investment and considering other opportunities. Some industries are thus more affected than others in times of financial "bubble" (Das, Jagannathan, and Sarin, 2003).

It can also be argued that when IPO markets are morose and pessimistic, i.e. bearish, venture capitalists refrain from investing in firms that are in a later development stage since the prospects of a profitable IPO are reduced. They may instead chose to divert their investments towards companies that are at an earlier stage, with a possible exit further away in time, in order to be sure that markets will have recovered when the time of going public comes.

The following measure of IPO activity is included in the study:

- IPO_MARKET_GLOBAL: gives the number of IPOs occurring during the year of the funding round on the most important US stock markets (AMEX, NYSE, and NASDAQ).
- IPO_MARKET_TECH: gives the number of tech IPOs occurring during the year of the funding round. Tech IPOs are defined here as IPOs

involving internet-related stocks plus other technology stocks including telecom, but not biotech.[5]

Table 1 gives the values that these variables can take.

IPO Markets	Global	Tech
2006	157	48
2007	159	75
2008	21	6
2009	41	14
2010	91	33
2011	81	36
2012	93	39
2013	157	43

Table 1: Number of IPOs in the global and tech markets

B. Round-related variables

Round-related variables are used to describe the financing round. They are fixed by the contract between the investors and the firm.

1. The number of participants

More often than not, venture capital deals are syndicated – syndication arises when venture capitalists invest jointly in projects – and the size of the syndicate increases with the complexity of the deal and the amount required to support the company's growth.

The rationale behind the syndication of venture capital deals has been researched. In addition to the benefits of risk diversification, Lerner (1994b) studies the advantages of syndication for the selection process of venture capitalists. He suggests that a venture capitalist, even after its own evaluation of an investment opportunity, might still be unsure about the venture's prospects and might prefer to get the opinion of another venture capitalist. Brander et al. (2002) evoke the added value of the complementary management skills of additional venture capitalists, risk sharing and project scale as possible benefits of syndication. Another aspect of the added value of syndication is that, as the number of participants increases, the pool of available contacts increases as well, making it easier to find strategic buyers. Moreover, syndication with more experienced venture capitalists adds a reputation effect and signal of the investment's quality, thus facilitating an exit through IPO.

[5] The data about the number of IPOs is based on the numbers from Jay Ritter, available at: https://site.warrington.ufl.edu/ritter/files/2016/03/Initial-Public-Offerings-Updated-Statistics-2016-03-08.pdf

It can therefore be expected that a larger syndicate will reduce the time until a trade sale or an IPO and delay any liquidation as well as increase the probability of a successful exit.

The PARTICIPANTS variable gives the number of participants in a given financing round.

2. The amount

The amount received by the firm is defined by the AMOUNT variable. This variable is expressed in $1000.

It has previously been mentioned that in some industries it is not surprising to see companies collecting significantly higher amounts of money at the same stage of development than in other industries. Nonetheless, no clear relationship has been established between the amount involved in the financing rounds of these industries and their exit strategy.

There are reasons to imagine that more funds supplied to the company should decrease the time until a favorable exit. For example, more money should give entrepreneurs more resources, thus facilitating the success of their projects and decreasing the time needed to complete them. On the other hand, it gives them more time to pursue unsuccessful projects, either until they start paying back or are abandoned, thus increasing the time to exit.

Previous survival analysis has found that the amount has a positive effect on the time to exit. This variable should behave in a similar fashion in this analysis.

3. The presence of business angel

ANGEL_DUMMY is a variable that is equal to 1 if there is at least one business angel among the investors and 0 otherwise.

The point of this variable is to examine the tangible consequences of the presence of business angel on the exit type and timing.

However, it is important to keep in mind that the database only references companies that have received venture capital money from either venture capitalists or business angels (or a mix of both). The objective is not to analyze the differences in the exit strategies between companies that have received money from business angels and any other companies, thus including those that have not received venture capital money. The conclusions reached by investigating this variable will therefore be limited to those firms that have received venture capital.

Previous literature on business angels widely recognizes that they tend to be present at a much earlier stage than venture capitalists. It will thus possible to see if their presence among the investors has a long-lasting effect on the company through the analysis of the exit strategy of firms that have an “angel_dummy” variable equal to 1 and the others.

It is not clear what to expect from this variable. On the one hand, it could be argued that the effects of the presence of business angels at an early stage might be overshadowed by venture capitalists’ presence at later stages, as successful business will encounter a successful exit, whether or not business angels are present. This would mean that there should be no distinguishable difference in the survival rates of the two groups. On the other hand, it could also be contended that it is the presence of business angels at an early stage that put the company on its successful path, towards later funding by venture capitalists and a favorable exit. In this case, the survival rates will be significantly different.

4. The presence of at least two venture capitalists

The variable VC_DUMMY is equal to 1 when at least two venture capitalists are present in the financing round and 0 otherwise.

By incorporating this variable, it becomes possible to differentiate the financing rounds that involve only business angels or only one venture capitalist in order to examine the specific impact of syndicated venture capitalist deals. The “participants” variable only gives the total number of participants and no distinction is made between angel investors and venture capitalists.

The same remarks as those made for the number of participants remain valid in this case if:

- There are no differences between the impact of venture capitalists and business angels on the exit strategy (this is also examined by the previous variable: “angel_dummy”).
- Or syndication of the venture capital deal has no impact on the exit strategy.

Furthermore, if either of these conditions is not met, then the survival rates of the two groups – when the variable is equal to 1 and when it is equal to 0 – will be significantly different. If this variable shows a lower survival rate for IPOs and trade sales when it is equal to 1, this would mean that the benefits of syndication on the time to exit are real: adding more venture capitalists to a deal effectively reduces the time until a favorable exit is achieved. It would then be preferable to have two or more venture capitalists than having only one, or none if only business angels are involved.

The opposite would be true if the survival rates of the groups are not distinguishable. It would then be concluded that the fact that there are only business angels involved or the fact that there are less than two venture capitalists makes no significant difference on the exit strategy compared to when there are two or more venture capitalists.

C. Variables of interest

The variables of interest are those that are being studied. These variables, obviously, are not known at the time of the investment.

1. The duration

The DURATION variable is calculated as the number of days between the date on which the round started and the time of the exit (if the investment was exited).

If the investment was not exited, the duration becomes the numbers of days between the date on which the investment round started and the date of collection of the data, in this case December 31, 2013. The duration of such rounds is called right-censored. Censoring arises when the event of interest occurs at a time outside the time interval of the study.

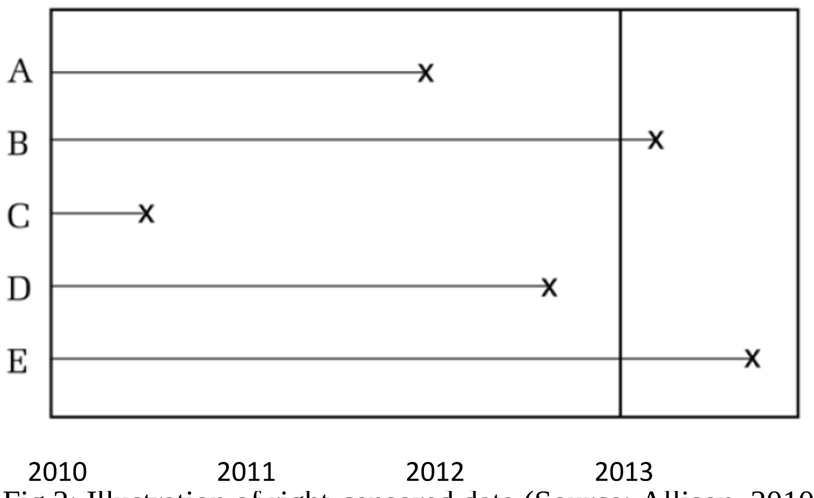


Fig 2: Illustration of right-censored data (Source: Allison, 2010)

Suppose that Fig 2 depicts some funding rounds that, for the sake simplicity, all started at the same time in 2010. The horizontal axis stands for the time. Each line labeled A through E represents a different round, in other words, the survival of a company until an exit occurs. A cross at the end of the line indicates that an exit occurred at that point in time. The line at 2013 represents the beginning of the study and the date of collection of the data. Any exit occurring in 2013 or earlier is observed, and hence these exit times are uncensored. Any exit occurring after 2013 is not observed, and these times are censored at time 2013. Therefore, rounds A, C, and D have uncensored times to event, while rounds B and E have right-censored times to event.

This variable is the main focus of the study since it represents the life of a company from a given round until the exit (or until December 31, 2013).

2. The exit type

The EXIT variable gives the status of the investment and can take several values:

- 1 if the investors exited through a trade sale (acquisition)
- 2 if the investors exited through a write-off (liquidation)
- 3 if the investors exited through an IPO
- 0 if the investors have not yet exited (at the time of the data collection) or have exited through another route.

It should be pointed out that the majority of the financing rounds are characterized by EXIT = 0.

Besides, if a company has more than one financing round, this variable takes the same value for all financing rounds since only the exit type that occurs at the end of the investment is considered (at the end of the last round).

All these expectations provide some first intuitive answers to the research questions and hypothesis as well as guidelines for the empirical study and the definition of the model. Of course, these expectations will have to be either confirmed or negated when analyzing the results of the model.

IV. Data

As previously mentioned, the data used in the empirical analysis was extracted from the freely available snapshot of the CrunchBase database. This snapshot is a replica of the CrunchBase dataset from December 2013.[6] The information is downloadable in the form of a MySQL dump, scattered in different MySQL tables, and some knowledge of this programming language is needed to gather every useful piece of information.

The final database use in this study consists of successive records ("spells" in the terminology of survival), each record corresponding to one investment round in a venture-backed firm. When the firm was involved in several financing rounds, there is a corresponding number of records in the database.

Although the CrunchBase snapshot contains 52.928 different financing rounds, I had to apply some selection criteria in order filter out the most interesting data.

First the financing rounds that occurred before 2006 were removed, that is to say approximately one year before the creation of the website. I considered that before this date, the financing rounds were referenced too inconsistently to be of use and would just represent unnecessary outliers for the analysis. Only around 3.000 rounds belong to the 1999 to 2005 period (compared to 49.892 for the 2006 to 2013 period). The rounds with missing data on the amount were removed as well (5.575). The rounds in which venture capitalists were not involved are of no interest and are therefore discarded (3330 debt or crowd funding rounds).

Then in order to emphasize on the strengths of the database, that is to say where is has the most chances of being complete and representative of the actual operations of venture capitalists, I remove all rounds that do not take place in the United States (12.506). For the same reason I also discard the rounds concerning companies that have no chance of being directly or indirectly related to the internet (i.e. biotechnologies, semiconductors, pharmaceuticals, clean-tech, industrial, transportation, and so on and so forth). Indeed, the major target of CrunchBase are the high-tech Internet startups, other industries are not referenced as extensively.

After filtering the data I obtain a dataset made up of 19.331 investment rounds for 11.499 distinct companies.

---

[6] Available at: https://data.crunchbase.com/docs/2013-snapshot

As previously mentioned, the CrunchBase dataset has not been a very popular source of data in the scientific literature. One reason for this might be that the website is not considered to be as reliable as other specialized data providers due to its information collection process, even though the information is verified before being edited into the site. Another explanation could be that it has a strong focus on Internet-related companies and does not reference other industries as extensively, which could lead to bias towards the web industry in empirical studies. Another possible objection is that it is too young to provide enough historical data for some studies.

Nonetheless, Block and Sandner (2009) used CrunchBase data on Internet startups for their paper. They investigated the coverage and representativeness of the database by comparing it with the industry statistics published by the National Venture Capital Association and found a significant correlation. In this paper also mostly focuses on – but are not limited to – Internet companies, since the database may not yet be representative of other industries.

The newness of the database is a potential concern. Meaning that most of the companies it references are also quite new, and thus have not yet been exited. For this reason, almost 80% of the final dataset is composed of still “operating” firms. There is no obvious solution to this problem.

One conceivable option would consist of simply discarding all, or most, of the censored observations. However, this would cause two problems. First, I would be removing some valuable information as well. Suppose that the study was about a rare disease, where finding participants was already difficult: I would then end up working with an even smaller dataset.

Second, this poses the problem of *informative censoring*. Imagine a study about a clinical trial, in which the duration information relating to subjects still alive at the end of the study is censored (right-censored). If a person who is responding well to treatment survives until the end of the study, this could be an indication that the person is doing well, thus indicating that they have a survival time longer than the censored time. By removing such information, not only would valuable data be discarded, but it would also be *systematically* discarded. More specifically, I would be systematically removing only the observations with higher survival times, and therefore biasing the data. While most survival analysis techniques work relatively well with censored data, all of them assume that the censoring is non-informative.

None of the books or academic papers that I have consulted offer solutions to the problem – or even bring up the topic – of highly censored data. Other studies on exit strategies that have used different databases have been done under lower levels of censoring. One possible reason for this is that professional financial data providers may focus more on providing information on the conditions of the exit, and less on referencing each and every financing round. For example, they may look at the different financing rounds retrospectively, after the company is acquired or goes public. While the objective of CrunchBase is to provide general information about the ecosystem of innovating companies and the actors involved, there is no particular focus on those companies that have exited.

In the same vein, due to the newness of the website, the maximum possible duration of investments is lower than in other studies. In this case, the maximum duration of an investment is approximately 3.000 days – slight over eight years – for an investment round starting in 2006 and not exited at the end of 2013. This still gives plenty of information to work with, and in any case, empirical evidence from the US and Canada shows that venture capitalists stay for four to six years on average with their investment before being able to exit them (Cumming and MacIntosh, 2003b). Other studies have found average exit times between 900 and 2.000 days depending on the type of exit. However, while it should not be a problem when estimating the model, this may limit the perspectives when describing the dataset with descriptive approaches.

V. Methodology

This section details the framework of the empirical study. As mentioned in the introduction I use survival analysis methods to analyze the dataset. I will thus start by briefly defining survival analysis and its advantages over other statistical methods. The reasons for this choice have already been briefly evoked, but I will describe those reasons in greater depth here and specify which methods and models are used and why. Then I will discuss how these models can be adapted to competing risks. But before approaching the competing risks models, I will first summarize some key concepts and basic models of survival analysis. The following explanations are based on Jenkins (2005) and Allison (2010).

A. What is survival analysis?

Survival analysis is a branch of statistics for studying the occurrence and timing of events. It was originally designed to study death, hence the name. However, these methods, can be used in a vast number of fields in natural sciences, sociology, engineering, economics, and so on.

Survival analysis methods were designed to analyze longitudinal data on the occurrence of events. To be studied through survival analysis, an event can be defined as a transition from one discrete state to another. Ideally, the change occurs instantaneously so that it is possible to determine precisely when it happened. A typical objective of survival analysis is to estimate causal or predictive models in which a set of covariates determines the risk of an event. The dataset must then contain measurements of such covariates in addition to the time of the events. The covariates can be fixed over time or time-dependent, leading to different specifications for the model. Once the model is established, one can, for example, estimate the time to event for a group of individuals, compare it between two groups, assess its relationship with regard to covariates, and so on.

B. Why use survival analysis?

The main advantage of survival analysis over other statistical methods resides in the fact that survival analysis can handle censored data. Observations are called censored when the information about their survival time is incomplete. Specifically, right-censoring occurs when an individual has not experienced the event of interest at the time of the study. The survival time for this individual is thus considered to be at least as long as the duration of the study.

Let us suppose that I estimate an ordinary linear regression model by setting the variable of interest – the dependent variable – as the time to exit. What could be done about the investments that had not been exited at the time of the study, i.e. the censored observations? One option would be to discard each censored record. This might work if the proportion of censored observations was small. However, such cases represent the large majority of the sample. In addition, the fact that an event has not occurred – yet – is not completely irrelevant. It is informative and should therefore be incorporated into the model. Besides, it has been shown that discarding all censored records may result in significant biases in the estimation of the parameters, because of the informative censoring problem discussed earlier. Another option would be to set an arbitrary duration for these censored observations (the median of the other observations, for example). This is a strong assumption concerning a large number of observations, and again, may result in significant biases.

Another possibility would be to build a model using logistic regressions. Logistic regressions use a binary dependent variable: either the investment is exited or it is not. But in this case, the analysis completely ignores the time factor. The duration cannot be the dependent variable since it has to be a binary variable, and it cannot be an explanatory variable without leading to significant bias (to incorporate the duration it would be necessary to, again, either discard or "corrupt" the censored observations).

By contrast, any survival analysis method allows the censored observation to be incorporated in the dataset. It is designed to estimate the covariates and to incorporate the time factor. It also enables the analysis of the interactions between each type of exit through a competing risks model (more on this later). Table 5.1 summarizes the different aspects of each method.

Method	Explanatory Variables (covariates)	Dependent Variable	Handle censoring?
Linear Regression	Categorical or Continuous	Continuous	No
Logistic Regression	Categorical or Continuous	Binary	No
Survival Analysis	Categorical or Continuous	Continuous (time)	Yes

Table 2: Regressions versus Survival Analysis

C. Introductory notes

In survival analysis, the survivor function will be the most commonly used. It is defined as the probability of surviving beyond a point in time *t*, i.e. the probability that the event of interest occurs after a specified period of time. The survival function is therefore a special kind of *cumulative distribution function*, *cdf*. The cdf of a random variable *T* (in this case, *T* represents the time to exit, the time that an investment will "survive" before being exited), noted *F(t)*, is defined as the probability that the random variable *T* will be less than or equal to any value *t* that is arbitrarily chosen. The cdf is thus noted: $F(\mathrm{t}) = \Pr\{\mathrm{T} \leq \mathrm{t}\}$. The survivor function, however, is the opposite since it gives the probability that T will be greater than some arbitrary point in time *t*; it is thus defined as: $S(\mathrm{t}) = \Pr(\mathrm{T} > \mathrm{t}) = 1 - F(\mathrm{t})$. Since it is a probability, $S(\mathrm{t})$ is comprised between 0 and 1, and since *T* represents a time lapse it cannot be negative. Moreover, $S(0) = 1$, i.e. the probability that an investment lasts longer than 0 days is certain, and as *t* increases, $S(\mathrm{t})$ approaches 0, i.e. the probability that an investment lasts forever becomes non-existent. In Chapter 6 I will provide some estimates of the survivor functions of the dataset.

When variables are continuous, their probability distribution can be described using the *probability density function*, *pdf*. This function gives the probability of an event occurring at exactly time *t* (out of all the possible values of *t*), and is defined as:

$$f(t) = \lim_{\Delta t \to o} \frac{\Pr(t < T \leq t + \Delta t)}{\Delta t} \quad \text{or equivalently} \quad f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt}$$

In the case of survival analysis, since individuals experience the event of interest only once (an individual can only die once, and an investment can only be exited once), they are no longer at risk of experiencing it again after it has occurred. This is why the *instantaneous risk* that an observation will experience the event of interest at time *t* is a conditional probability, simply because it first has to survive up to time *t*. Furthermore, since time is continuous, the probability that an event occurs exactly at time *t* is necessarily 0. But the probability that the event occurs in a small interval of time could instead be considered, say between *t* and $t + \Delta t$. This describes the *hazard function*, written as

$$h(t) = \lim_{\Delta t \to o} \frac{\Pr(t < T \leq t + \Delta t \mid T \geq t)}{\Delta t}$$

This instantaneous risk, called the *hazard rate*, given by the hazard function, is defined as the instantaneous risk of failure for the individuals that have survived up to time *t* to experience the event of interest during the next instant of time. The hazard function is preferred in survival analysis over the pdf because its attempt to quantify the instantaneous risk that an event will take place at time *t* is conditioned by the survival of the object up to time *t*. The hazard function is always positive and when *h*(t) = 0, it implies that the event cannot occur at time *t*.

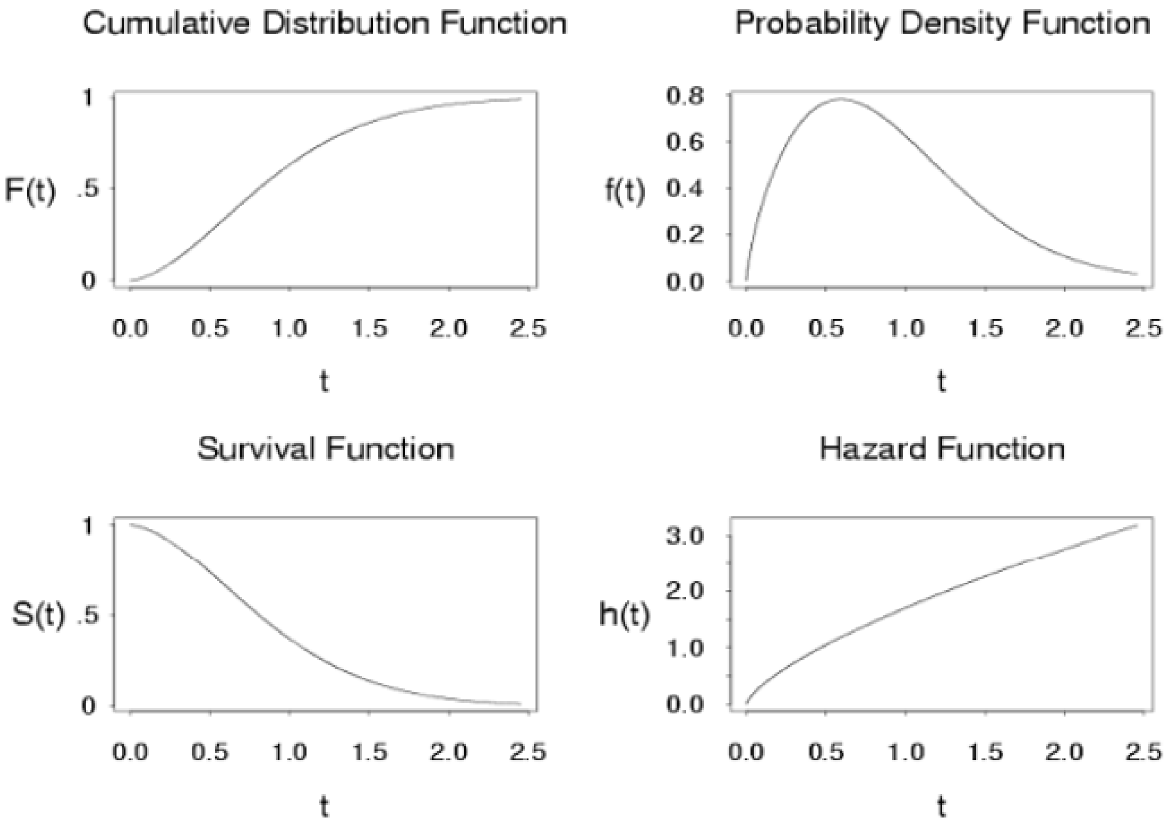


Fig 3: The different survival analysis functions

As shown in Fig 3, the survivor function and the hazard rate are two sides of the same coin: when the survivor function decreases more (less) sharply in an interval, it means that the hazard rate for this interval is higher (lower). They are linked to the probability density function by the following formula:

$$h(t) = \frac{f(t)}{S(t)}$$

Before proceeding, note that while the hazard seems to have an easy interpretation as an instantaneous probability, it is not a probability because it can be greater than 1 (but cannot be negative): this may happen because of the division by Δt.

More specifically, the hazard is a quantity that can be defined as the number of events per interval of time (hence the reason why it is sometimes called hazard *rate*), just as a speed is measured in kilometers per hour. Suppose that the hazard for earthquakes in a region is 2.48 per year, this means that assuming a constant hazard for the year, the region is expected to experience 2.48 earthquakes over the course of the year. But what about events that are not repeatable, such as death or the exit of an investment? By taking the reciprocal of the hazard it is the expected time until the occurrence of the event that is obtained. Suppose that the death hazard of an individual at a point in time is 0.023 per year, then the individual can expect to live 1/0.023 = 43.5 more years. This is assuming that the individual and his or her environment does not change. Similarly, the speed of a car going at 100 kilometers per hour can be interpreted as meaning that if the car followed its trajectory at this exact speed for one hour, it would travel 100 kilometers. In reality, however, the hazard will most certainly fluctuate as the life of the individual continues, for example, as the individual becomes older, his or her hazard will probably increase.

Regarding investments, suppose that every venture-backed company carries a hazard for every kind of exit, i.e. a hazard for IPOs, a hazard for trade sales, and a hazard for liquidations. As time goes by, if the company is successful, its hazards for IPOs and trade sales increase while its hazard for liquidations decreases. If the company does not perform as well as expected, it would assumed that its hazard for IPOs would start to decline, and after some time, if it continues to disappoint, its hazard for trade sales will decline as well, while its hazard for liquidations will increase even more. This highlights the fact that the actual hazard function varies significantly with the surrounding environment. In fact, it is a function that is typically characterized by sharp increases or decreases, as the underlying observation moves from one situation to another. In the case of venture capital investment, this has led authors such as Giot and Schwienbacher (2005) to model exit time using non-monotonic functions[7] such as the generalized gamma distribution.

This is also a reason why hazards can be seen as characteristics of individuals instead of populations or samples (unless each individual in the population is the same) and why different individuals can have completely different hazard functions. When modeling the hazard function of a sample, the individual specificities tend to be smoothed out to reveal the much larger trends, the bigger picture.

[7] A monotonic function is a function that is either entirely increasing or decreasing.

D. The accelerated failure time model

Typically, the accelerated failure time model (AFT) describes a relationship between the survivor function of two individuals, and assumes that the effect of covariates is to accelerate or decelerates the life of an individual at risk of the event of interest, in other words, that what differentiate two individuals is their aging rate. By contrast, the proportional hazards model, the other commonly used model (see next section), assumes that the effect of a covariate is to multiply the hazard by some constant.

Let us take the survivor function of two individuals denoted by S_1(t) and S_2(t). The AFT models states that there is a constant Φ > 0, specific to the combination (1; 2), such that

$$S_1(t) = S_2(\Phi t) \text{ for all } t \geq 0 \tag{5.4.1}$$

This suggests that individual 1 will age at a rate that is Φ times the rate of individual 2. To illustrate, suppose that $S_1(t)$ stands for the population of humans and $S_2(t)$ the population of dogs. The popular wisdom says that a year for a human is equivalent to seven years for a dog, this implies that Φ = 7 and $S_1(t) = S_2(7t)$. So the probability of a human surviving 70 years or more is the same as the probability of a dog surviving ten years.

Now let us imagine a sample of *n* individuals, described by a set of *k* covariates. Let T_i be a random variable denoting the – possibly censored – time to event of individual *i* and let $x_{i1}, x_{i2}, \dots, x_{ik}$ be the measurement of the *k* covariates of individual *i*. According to, among others, Jenkins (2005), Zhang (2005) and Allison (2010), the classic corresponding AFT model is:

$$\log(T_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \sigma\varepsilon_i \tag{5.4.2}$$

Equation (5.4.2) can be written equivalently, by exponentiating each side of the equation, as

$$T_i = \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \sigma\varepsilon_i) \tag{5.4.3}$$

where β_0, …, β_k are the regression coefficients corresponding to each covariate, σ is a scale parameter, and ε_i stands for the random disturbance term. Equation (5.4.2) is very similar to a standard linear regression – such as equation (1.1) – the only differences being the presence of a parameter σ before the disturbance term ε and the fact that the dependent variable T_i is logged (the reason why T_i is logged is straightforward considering that time to event is always positive).

It was already mentioned that in a linear regression model, the random error term e_i is usually assumed to be an independent and identically distributed (i.i.d.) random variable following a normal distribution $N(0; \sigma^2)$. It follows that e_i can be denoted as $e_i = \sigma\varepsilon_i$, in which case the ε_i's are i.i.d. from $N(0; 1)$. In other words, the parameter σ can be ignored, but this requires that ε_i can vary from one observation to another. However, it is more practical to fix the variance of ε at some standard value – i.e. 1 – and let the value of σ change to accommodate changes in the disturbance variance (Allison, 2010).

The regression coefficients – $\beta_1, \ldots, \beta_k$ – in equation (5.4.2) have an interpretation very similar to those of a linear regression. When $\beta_k > 0$ ($\beta_k < 0$) it can be interpreted as the percentage increase (decrease) in the average survival time when the value of the covariate x_k is increased by one unit. Therefore, the greater the value of the covariate corresponding to a positive β_k, the more the survival time of the observation is prolonged. Conversely, an increase in the covariate value corresponding to a β_k negative leads to a reduction of the expected survival time; the life of the observation has been "accelerated". The estimated β coefficients determine the "time acceleration" for the corresponding observations.

Furthermore, taking the exponential of the coefficient gives the *time ratio*, the estimated ratio of the expected mean survival times for the two groups. For example, if $\beta = 0.58$ then $e^{0.58} = 1.79$ which means that, all other things remaining equal, the expected time to event for one group is 79 percent greater than the expected time to event of the other group. Moreover, dividing the exponentiated regression coefficient of one population by that of another population gives relative time ratios. For example, the relative time ratio of a population of men against a population of women is $e^{\beta_{Men}}/e^{\beta_{Women}}$. For quantitative covariates, the transformation $100(e^{\beta} - 1)$ gives the percent change in the expected time to event if the corresponding covariate value increases by 1 unit.

The AFT model is parametric, requiring the specification of a distribution for the disturbance term ε_i (equivalently for T_i). For example, assume that ε_i are i.i.d. following a normal distribution $N(0; 1)$ this is equivalent to assuming that T_i follows a log-normal distribution. As shown by Table 3, several distributions for T_i are possible, giving their name to the corresponding AFT models:

Distribution of ε	**Distribution of *T***
extreme values (1 parameter)	Weibull
extreme values (2 parameters)	exponential
log-gamma	gamma
logistic	log-logistic
normal	log-normal

Table 3: Possible distributions in AFT models

The choice of a distribution can be influenced by the hazard function to be modeled. For example, Giot and Schwienbacher (2005) have suggested the use of non-monotonic functions, and particularly the generalized gamma distribution, in order to model the time to exit of venture capitalists.

Others, such as Cleves et al. (2004) or Allison (2010), use fit statistics, for example the Akaike Information Criterion (AIC), to help determine the best suited model and baseline distribution. However, these statistics does not constitute a formal hypothesis test, so the comparison is only informational.

In this case, the AIC test for the generalized gamma distribution leads to lower values than other distributions, indicating a better fit for the model.

I have therefore selected the generalized gamma distribution. The generalized gamma model is specified as

$$f(t) = |\delta| \left(\frac{t^{\delta}}{\delta^2}\right)^{\frac{1}{\delta^2}} \exp(\frac{-\frac{t^{\delta}}{\delta^2}}{t\Gamma\left(\frac{1}{\delta^2}\right)})$$

where δ is called the "shape parameter". The generalized gamma distribution is fitted to the data assuming that $T_0 = e^{\varepsilon}$ (Zhang, 2005).

Fig 4 represents some possible hazard functions for the generalized gamma distribution (not the hazard function for the standard gamma distribution).

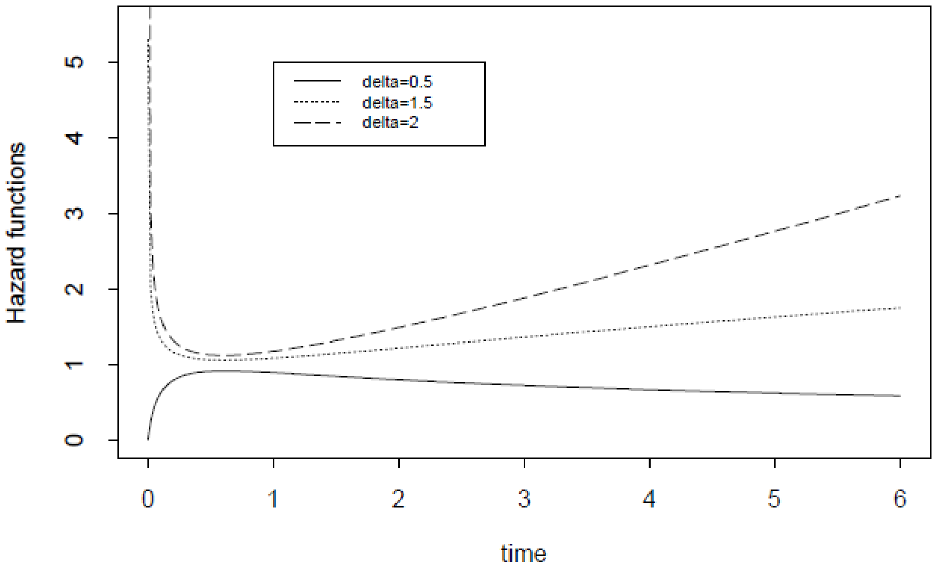


Fig 4: Possible hazard functions for the generalized gamma distribution

An advantage of the generalized gamma distribution is that it can take forms unlike any other distribution. For example, it is clear from this plot that when $\delta < 1$ the hazard function takes the form of an inverted U-shape, while when $\delta > 1$ the hazard function takes a U-shape. In addition, the generalized gamma distribution includes many other distributions, such as the exponential, Weibull, and log-normal models. This is why this distribution is one of the most flexible for survival analysis. On the other hand, the generalized gamma distribution cannot represent hazard functions characterized by more than one reversal of direction.

The AFT models are estimated using the *maximum likelihood* method. This method produces estimators that have good large-sample properties. Under certain conditions, maximum likelihood estimators are consistent, asymptotically efficient, and asymptotically normal. Consistency means that, as the sample increases, the estimates converge in probability to the actual values, meaning that in large samples the estimates will be approximately unbiased. “Asymptotically efficient” means that, in large samples, the estimates will have standard errors that are at least as small as those for any other estimation method. And, finally, “asymptotically” normal means that the sampling distribution of the estimates will be approximately normal in large samples, which implies that the normal and chi-square distributions can be used to compute confidence intervals and p-values. The mathematics behind the maximum likelihood method are beyond the scope of this discussion, for more details refer to Kalbfleisch and Prentice (2002) for example.

E. The Cox proportional hazard model

First introduced by Cox (1972), the most commonly used model is the Cox proportional hazard model (PH model), which focuses directly on the hazard function. The typical feature of Cox's proportional hazard model is that it makes it possible to estimate the relationship between the hazard rate and the covariates without having to make any assumptions about the shape of the baseline hazard function. This is why it is called a semi-parametric model, as opposed to the parametric models considered in the previous section. The general model states that the hazard for an individual i at time t with k covariates x_{ik} is usually written as

$$h_i(t) = \lambda_0(t)\exp(\beta_1 x_{i1} + \cdots + \beta_k x_{ik}) \tag{5.5.1}$$

According to this equation, the hazard is therefore composed of two multiplying components:

- An exponentiated linear function of the covariates:

  $\exp\,(\beta_1 x_{i1} + \cdots + \beta_k x_{ik})$
- A function $\lambda_0(t)$ that is left unspecified.

In fact, $\lambda_0(t)$ corresponds to the hazard function when the values of all the covariates of an individual are 0. It is called the *baseline hazard function*. This baseline hazard function serves as a reference while the other component is the relative risk (a proportionate increase or reduction in risk) associated with the set of covariates of an individual. Let us imagine that two groups differ only by one covariate *x*, and that the two values this covariate can take are 1 (for group one) or 0 (for group zero); the model becomes:

$$h_i(t) = \begin{cases} \lambda_0(t) & if\ x = 0 \\ \lambda_0(t)\exp(\beta) & if\ x = 1 \end{cases}$$

Thus, $\lambda_0(t)$ is the risk at time *t* for individuals in group zero, and $\mu = \exp(\beta)$ the risk in group one relative to group zero at any time *t*. Then, if μ takes the value 1 (or equivalently $\beta = 0$) the risks are the same in the two groups. But if $\mu = 7$ (or $\beta = 1{,}9459$), then the risk for individuals in group one at any given time is 7 times the risk of individuals in group zero with the same age. Therefore, the *hazard ratio* of group one compared to group zero is 7.

The hazard ratio (HR) is thus the ratio of the hazard functions of two populations differing by the two levels of a covariate, and can be interpreted as the chance of an event occurring in one group divided by the chance of the event occurring in the other group. A hazard ratio of 1 means that the hazard functions of the two groups are similar, whereas a value other than 1 indicates a difference in hazard rates between the groups. The hazard ratio between two groups can be linked to the survivor function of the corresponding groups through the following formula:

$$S_0(t) = S_1(t)^{\mu}$$

where S_0 and S_1 are the survivor functions of group zero and one respectively. In the previous example, imagine that half the population of group zero survives until time *t*. The survival rate of group zero will be $S_0(t) = 0.5$, with a hazard ratio of 7, and the survival rate in group one becomes $S_1(t) = 0.5^7 = 0.008$ (0.8% of the individuals from group one survive until time *t*).

What is important when taking the ratio of the hazards for two populations is that the $\lambda_0(t)$ term cancels out of the numerator and denominator (the two groups have the same baseline hazard function, and the differences will emerge from the values of the covariates). As a result, the ratio of the hazards does not convey information about how soon the event of interest will occur: it is constant over time. This is because, while the assumed relationship between the hazard function and the covariates is not linear, the hazard ratio comparing any two groups is constant (all other things remaining the same). This is the assumption of *proportional hazard* of Cox's model. If the graphs of the log hazard function for any two individuals are plotted, the proportional hazard assumption implies that the hazard functions should be strictly parallel, as shown in Fig 5.

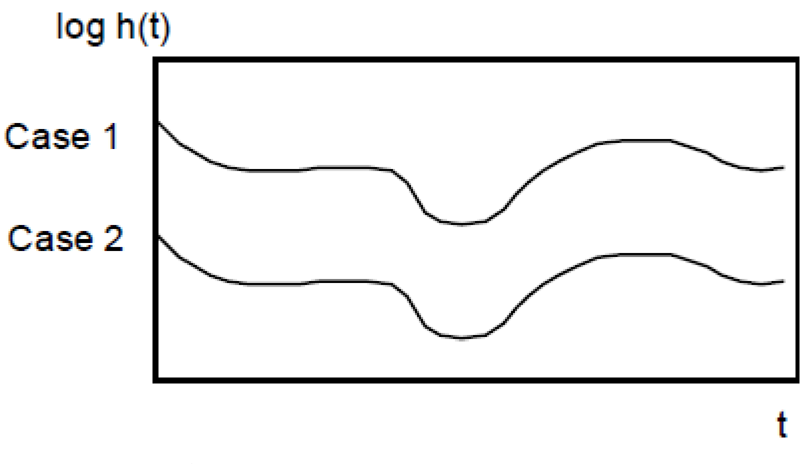


Fig 5: Log-hazard function in a Cox PH model (Source: Allison, 2010)

The fact that the baseline hazard function is completely unspecified may be a great advantage (i.e. by avoiding problems related to specifying the wrong distribution), but it can also be seen as a disadvantage, for example when researching the shape of the baseline hazard function.

This model uses the *partial likelihood* method to estimate the regression coefficient, as opposed to the AFT model which uses the maximum likelihood method. The partial likelihood estimates retain some properties of the maximum likelihood estimates: in large samples they are approximately unbiased and their sampling distribution is approximately normal. The asymptotically efficient property is lost.

The proportional hazard assumption is strong and can lead to significantly biased coefficient estimates when violated. Some authors therefore consider it to be often unreasonable (Cantor, 2003). Methods have thus been developed for testing and modeling non-proportional hazards. For example, in the presence of time-varying covariates the PH assumption no longer holds, because the time-dependent covariates will change at different rates for different individuals, so the ratios of their hazards cannot remain constant. One solution could be to subdivide time into intervals and assume that the baseline hazard is constant in each interval, leading to the piece-wise exponential model. This discussion, however, is beyond the scope of this paper (for more details refer to Alison, 2010).

The proportional hazard assumption has been tested for the dataset using the method proposed by Lin, Wei, and Ying (1993) based on the martingale residuals of the regression. The results show that several covariates significantly violated the assumption. Fig 6 depicts one such covariate. The dashed lines represent empirical scores based on 20 random simulations respecting the PH assumption. The solid line stands for the observed process. If it deviates significantly from the simulated processes, this is evidence that the PH assumption is violated for this covariate. For the variable VC_DUMMY, the observed process appears more extreme than the simulated paths. In the corner, the *p*-value (based on a Kolmogorov-Smirnov type supremum test) gives the percentage of simulated processes that had extreme points for 1000 simulated paths. For this variable, only 0.3% of the 1000 simulated processes had an absolute maximum exceeding that of the observed process. The assumption of proportional hazard must therefore be rejected for this variable.

Appendix 1 gives the results of the test for all covariates. These results indicate that several covariates strongly reject the assumption of proportional hazard.

This is one of the reasons why I preferred the AFT model over the Cox PH model.

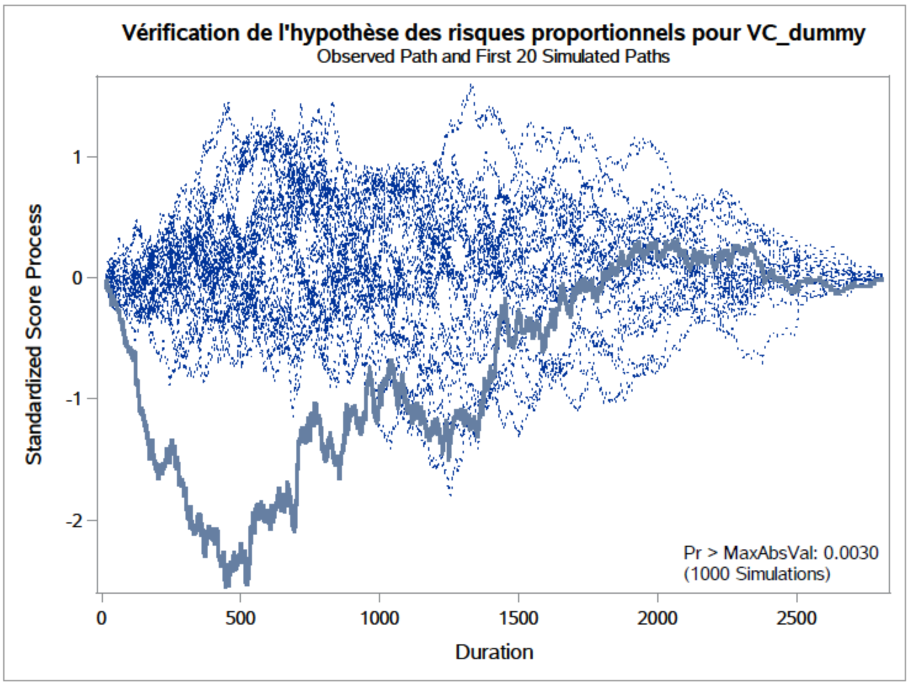


Fig 6: Proportional Hazard Assumption test for variable VC_DUMMY

Another reason is that most of the research on venture capital exit has been done specifying AFT models, and using the same model facilitates the comparison of the results. The interpretation of the coefficients is different for the Cox PH and AFT models. In a PH model, each regression coefficient indicates the proportional effect on the hazard rate of changes in the covariate, while the AFT regression coefficient indicates the proportionate change in survival time that corresponds to a change in the independent variable. The dependent variable in the Cox model measures the risk of experiencing the event of interest ($h_i(t)$ on the left-hand side of equation (5.5.1)), whereas in the AFT model it measures the survival time (T_i on the left-hand side of equation (5.4.3)).

The final reason is that, as previously mentioned, time ratios carry no information about the time factor of the risk of failure which is a key component of the study of venture capital exit. However, I estimated a Cox PH model to compare the consistency of the results obtained through the AFT specifications, these results will be discussed in Chapter 6.

F. Competing risks

So far, none of the models considered has made any distinction in the type of exit of the observation. Either the event of interest has been observed by the end of the study, or it has not. For example, in a clinical study that examines the death of patients who have undergone major surgery, this would be equivalent to knowing whether the patient is still alive at the end of the monitoring period, and if it is not the case, then knowing when the death occurred.

The cause of the event is therefore irrelevant to the analysis. But what about those who died of causes completely unrelated to the surgery? It seems essential to distinguish between those patients that die of complications due the surgery and those that die of, for example, a car accident or cancer. And in the case of venture capitalists' exit strategy, only knowing if an investment has been exited is of little interest if there is no information about, at least, the success of the exit.

Competing risks arise when an individual is at risk of more than one mutually exclusive event and the occurrence of one type of event removes the individual from risk of all the other event types. For example, venture capitalists do not have the possibility to exit their investment by both liquidation and trade sale, and once either of these exits has occurred the investment is no longer at risk of the other type of exit. For each type of exit, a separate hazard function has to be defined, a *cause-specific hazard function*.

As defined previously, T_i is a random variable that stands for the time of exit for round *i*. Let us now define another random variable J_i that denotes the type of the exit for round *i*. Thus, $J_i = 1$ means that round *i* is exited by trade sale, $J_i = 2$ means that the round is exited by liquidation, and $J_i = 3$ that it is exited by IPO. The cause-specific hazard function are defined as follow:

$$h_{ij}(t) = \lim_{\Delta t \to o} \frac{\Pr(t < T_i \leq t + \Delta t, J_i = \mathrm{j} \mid T_i \geq t)}{\Delta t}, \quad j = 1, 2, 3$$

The only difference between this definition and the definition of the hazard function is the appearance of $J_i = \mathrm{j}$. Thus, this calculates the conditional probability that a round is exited in the interval $[t, t + \Delta t)$ and that the cause of exit is of type *j*, given that the round was still "active" (not yet exited) just before time *t*. When one divides by Δt, this probability becomes a rate and the limit is then taken as $\Delta t \to 0$.

By the law of total probability, the general hazard of exit is the sum of all the cause specific-hazards (because exit must be due to one and only one of the causes)

$$h_i(t) = \sum_j h_{ij}(t)$$

The interpretation of cause-specific hazards remains the same. Any of the previously considered models can be adapted to incorporate dependence with respect to the covariates while factoring competing risks. The AFT model with three possible exits and k covariates is thus specified as follows:

$$\log(T_{ij}) = \beta_{0,j} + \beta_{1,j}x_{i1} + \beta_{2,j}x_{i2} + \cdots + \beta_{k,j}x_{ik} + \sigma\varepsilon_{i,j}\ , \quad j = 1, 2, 3$$

where the regression coefficients β are subscripted to indicate that the impact of the covariate depends on the exit type.

In the framework of exit strategies and with the explanatory variables defined in Chapter 3, I define our model as follows:

$$\begin{aligned}\log(T_{i,TS}) = {} & \beta_{0,TS} + \beta_{1,TS}ROUNDNB_i + \beta_{2,TS}\mathrm{MILESTONES\_ROUND}_i \\ & + \beta_{3,TS}\mathrm{IPO\_MARKET\_GLOBAL}_i + \beta_{4,TS}\mathrm{IPO\_MARKET\_TECH}_i \\ & + \beta_{5,TS}\mathrm{PARTICIPANTS}_i + \beta_{6,TS}\mathrm{AMOUNT}_i + \beta_{7,TS}\mathrm{ANGEL\_DUMMY}_i \\ & + \beta_{8,TS}\mathrm{VC\_DUMMY}_i + industry\ dummies + stage\ dummies + \sigma\varepsilon_{i,TS}\ ,\end{aligned}$$

$$\begin{aligned}\log(T_{i,LIQ}) = {} & \beta_{0,LIQ} + \beta_{1,LIQ}ROUNDNB_i + \beta_{2,LIQ}\mathrm{MILESTONES\_ROUND}_i \\ & + \beta_{3,LIQ}\mathrm{IPO\_MARKET\_GLOBAL}_i + \beta_{4,LIQ}\mathrm{IPO\_MARKET\_TECH}_i \\ & + \beta_{5,LIQ}\mathrm{PARTICIPANTS}_i + \beta_{6,LIQ}\mathrm{AMOUNT}_i + \beta_{7,LIQ}\mathrm{ANGEL\_DUMMY}_i \\ & + \beta_{8,LIQ}\mathrm{VC\_DUMMY}_i + industry\ dummies + stage\ dummies + \sigma\varepsilon_{i,LIQ}\ ,\end{aligned}$$

And

$$\begin{aligned}\log(T_{i,IPO}) = {} & \beta_{0,IPO} + \beta_{1,LIQ}ROUNDNB_i + \beta_{2,IPO}\mathrm{MILESTONES\_ROUND}_i \\ & + \beta_{3,IPO}\mathrm{IPO\_MARKET\_GLOBAL}_i + \beta_{4,IPO}\mathrm{IPO\_MARKET\_TECH}_i \\ & + \beta_{5,IPO}\mathrm{PARTICIPANTS}_i + \beta_{6,IPO}\mathrm{AMOUNT}_i + \beta_{7,IPO}\mathrm{ANGEL\_DUMMY}_i \\ & + \beta_{8,IPO}\mathrm{VC\_DUMMY}_i + industry\ dummies + stage\ dummies + \sigma\varepsilon_{i,IPO}\end{aligned}$$

Interestingly, each of the possible exits could be modeled using a different specification. For example, it would be possible to specify a log-normal model for the trade sale, a log-logistic for the liquidation, and a proportional hazard model for the IPOs. What makes this possible is that each model may be estimated separately for each event of interest with no loss of statistical precision. More specifically, because it is assumed that the occurrence of one type of event prevents the observation from experiencing any other types of event, such an observation therefore no longer contributes to the successive risk set (Lee and Wang, 2003). This implies that the likelihood function for each event type consider observations experiencing any other type of event as being censored at the time when the competing event occurred. The likelihood function of the model encompassing all the events can thus be factored into distinct likelihood functions for each event type taken separately (for more details regarding the likelihood functions of competing risks models, refer to Kalbfleisch and Prentice (2002) and Lee and Wang, 2003).

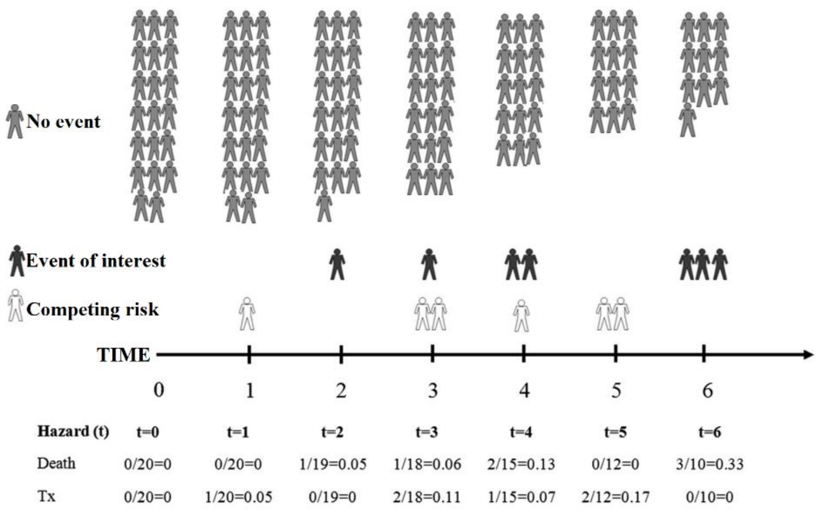


Fig 7: The cause specific approach (source: Lau et al., 2009)

Fig 7 depicts the calculation of the cause-specific hazard. The risk set starts at time *t* with 20 individuals (grey). Over time, the individuals experience either the event of interest (black) or the competing event (white). As events occur, the individuals are removed from the remaining risk sets. The calculation for the cause-specific hazard for both events is given at the bottom of the figure.

A second approach exists to in order to examine the effect of covariates in the context of competing risks called the *cumulative incidence function* (CIF). While the first approach, which consisted in applying typical survival analysis models to the cause-specific hazards, is more structural, focusing on the covariates of the risk of each type of event. This second approach is more descriptive, focusing on the probability of each event type.

An example of the difference between the two approaches is when investigating the effect of covariates. A covariate may appear to increase the occurrence of some type of events simply by lowering the rate of occurrence of events of other types, even if it has no effect on the rate of occurrence of the event in question.

The cumulative incidence function is defined as follows:

$$I_j(t) = \Pr(T_j < t, J_i = j)$$

This equation estimates the probability for an individual i of experiencing the event j before time t, in the presence of competing risks. A feature of the CIF is that at any time, the sum of the probability for each event plus the probability of no event is equal to 1. Graphically, this is simply represented by a step function that increments every time an event of type j occurs. Furthermore, Gray (1988) proposed a modified Chi-square test approach to testing the difference in CIF among two different groups.

The CIF method is of particular interest as a descriptive device. However, if the objective of the study is to examine the causality relationships between the event occurrence and the covariates, then the estimation of hazard specific function – by censoring the other event types – should be preferred (Pintilie, 2006).

This concludes the chapter dedicated to the key concepts of survival analysis, and detailing the reasons behind the choice of model and its possible specifications. The next chapter is devoted to the analysis of the results given by our model.

VI. Results

Before analyzing the results of the model, I will perform a preliminary examination of the dataset using descriptive methods. Then I will analyze the estimation results of the main model and use them to answer the initial research questions. And finally, I will compare the results of the model with other techniques to assess their robustness, and then go over the different methods I used to evaluate the goodness of fit of the model.

A. Descriptive analysis

The first thing that could be done is to estimate the survivor function of our sample for each type of event. To do this, the most widely used method is the Kaplan-Meier method.

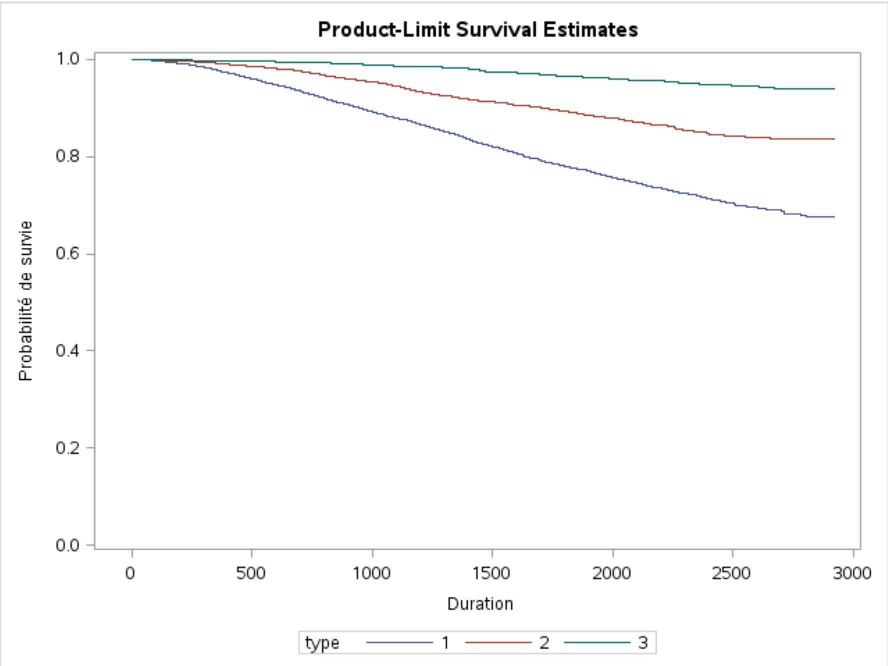


Fig 8: Plot of the survivor functions

The survivor functions are depicted in Fig 8. The variable TYPE determines the type of exit, and its values stand for, acquisition (1), liquidation (2), and IPO (3). The time is measured in days by the variable DURATION.

It is clear from the examination of this graph that the survivor functions are distinct. In addition, six tests of equality (namely Log-rank, Wilcoxon, Tarone, Peto, Modified Peto, and Fleming) strongly reject the null hypothesis that the groups have exactly the same survivor function; see Appendix 2 for the exact numbers.

Furthermore, it appears that the exit occurring the most often is acquisition, followed by liquidation and finally IPO. The "risk" for a company of undergoing an acquisition is the highest among the three possible exits. It also seems that the probability of companies "surviving" an acquisition starts decreasing sooner than for the other exits. For example, after only 500 days, the number of recorded acquisitions is much higher than the liquidations or IPOs. However, a common factor for all functions is that after some time, around 2500 days, the probabilities of exit appear to stabilize and reach a plateau.

Using the same method, the effect of covariates on the survivor function can be tested. For example, it is possible to graphically demonstrate the impact of the presence of business angels. This is represented by Fig 9.

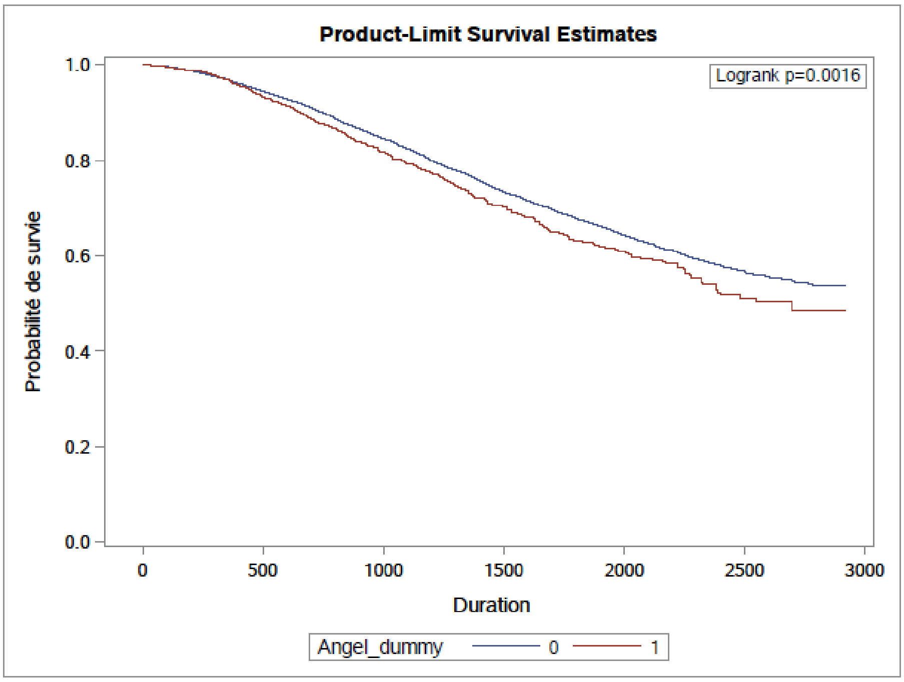


Fig 9: Testing the effect of business angels on survivor function

It appears that when business angels are present (when the variable ANGEL_DUMMY =1, the red line), the survival probabilities are lower, in other words, the probability of an exit are increased when business angels are present. Once again, all equality tests reject the null hypothesis of equality between each function. Unfortunately at this point it is not possible to differentiate between the types of exit. This means that since the probability of any exit is increased, it could be that some types of exit are more influenced by this variable than others. And if an increased exit rate for acquisitions or IPOs may be a good thing, it is not the case for liquidations. What can be done, however, is to only take into account exits by liquidation, and see if business angels make a difference to this precise type of exit. This is what is represented in Fig 10 and Fig 11.

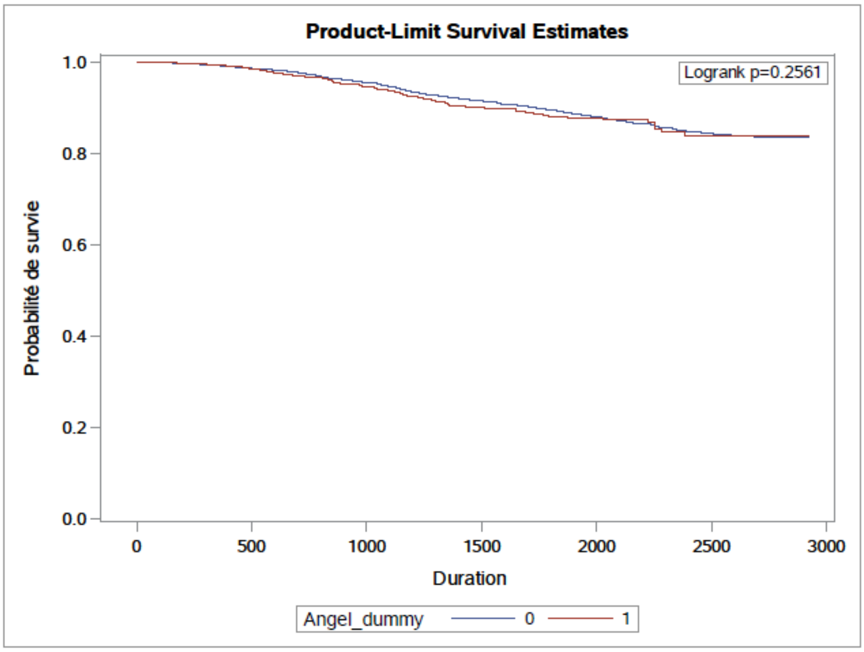


Fig 10: Testing the effect of business angels on exit by liquidation

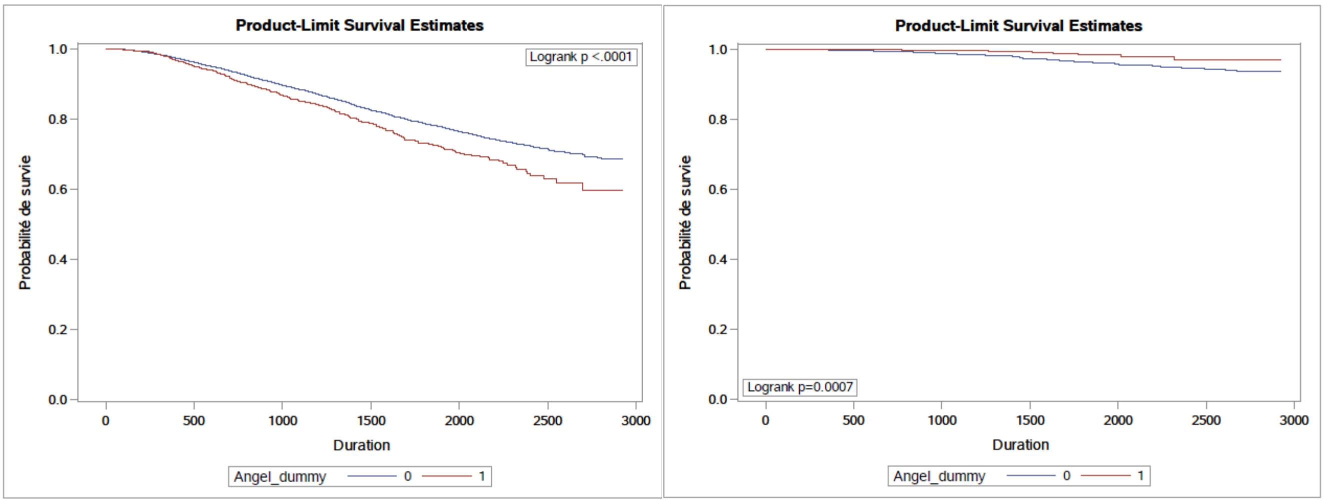


Fig 11: Testing the effect of business angels on exit by acquisition (left) and IPO (right)

In the case of exits by liquidation however, the two functions are nearly indistinguishable, as shown by Fig 10. This is supported by the high p-value of hypothesis test. Thus, there is no statistical evidence that would support the rejection of the null hypothesis that business angels have a – statistically significant – impact on the probability of occurrence of liquidations. It must be emphasized that this does not mean that business angels have no impact on the probabilities of liquidation, but that there is no statistical evidence in the dataset to conclude that there is a statistically significant difference between the two groups. For the other two exits however, this hypothesis is strongly rejected.

For exits by acquisition, the red line corresponding to the group of investment rounds that included business angels is below the line of the other group, meaning that the presence of angels increases the probability of this type of exit for the firm. But the opposite is the case for IPOs. I will go into more details about the covariates' impact on the survival of investments when discussing the results of the model. However, a first possible reason for this is that angels tend to invest earlier in the life of a company, for this reason their impact may be overshadowed by that of venture capitalists.

Now that the survivor functions have been examined, a question worth considering is whether the cause-specific hazard functions are proportional. More precisely, to determine if when one of the cause specific-hazard functions varies with time, another varies in a proportionate amount. This hypothesis can be formulated as follows

$$h_j(t) = \delta_j h(t), \;\; j = 1, 2, 3 \tag{6.1.1}$$

where the δ_j's denote some proportionality constant.

Fig 12 depicts the graphical examination of this hypothesis by plotting the kernel smoothed hazard function. This graph is a sort of moving average of the hazard functions.

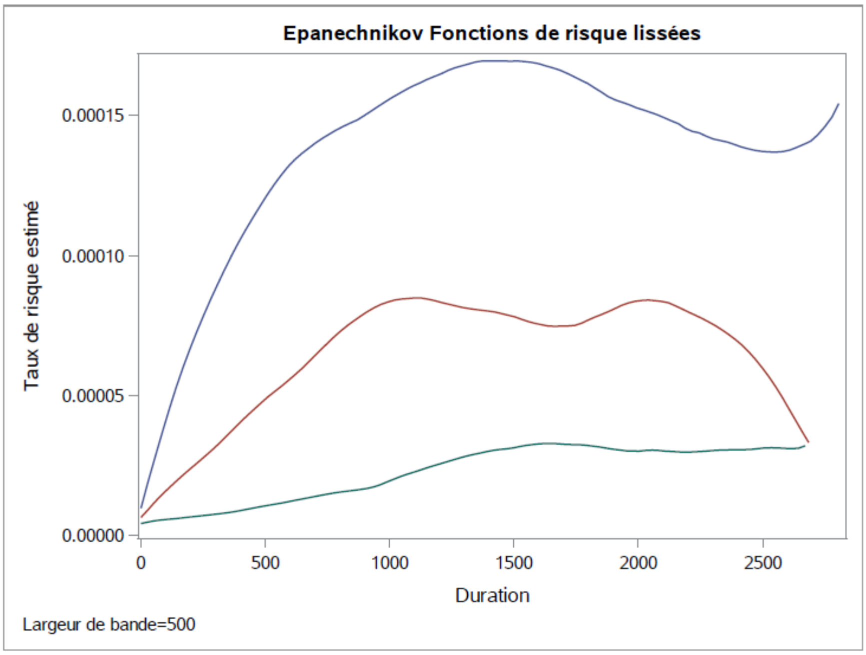


Fig 12: Kernel Smoothed Hazard Functions for all Types of Exits

This graph confirms that the risk of IPO is lower than the risk of liquidation, which is in turn lower than the risk of acquisition. There is also a tendency for the curves to move apart as time progresses for the first 1.000 to 1.500 days; this is evidence against the proportionality hypothesis. At 1.500 days, all risks seem to reach a peak, a bit earlier for liquidation. Towards the end of the analysis period, while the hazard function for IPO remains fairly constant, for liquidation it plummets, and for acquisition it shows signs of increase.

A parametric test proposed by Cox and Oakes (1984) can be used to evaluate the proportionality hypothesis in equation (6.1.1). Their test consists of a multinomial logit model for the type of event, with the time of the event included in the form of an explanatory variable. If the proportionality hypothesis is verified, the coefficient for the time factor should be 0. The results of this test are presented in the following tables.

Estimations par l'analyse du maximum de vraisemblance

Paramètre	Exit	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > Khi-2
Intercept	1	1	2.5973	0.1280	411.4658	<.0001
Intercept	2	1	1.6181	0.1359	141.7929	<.0001
Duration	1	1	-0.00051	0.000102	24.8014	<.0001
Duration	2	1	-0.00032	0.000109	8.9170	0.0028

Analyse des effets Type 3

Effet	DDL	Khi-2 de Wald	Pr > Khi-2
Duration	2	28.2732	<.0001

Table 4: Testing the proportionality hypothesis

First, by looking at the Type 3 table (on the right), it can concluded that the impact of DURATION is highly significant (p-value almost null), implying that the proportionality hypothesis is not verified. Next, the regression coefficient in the left table gives information about which hazard function might be proportional. The first row with parameter “Duration” is the contrast between the type 1 hazard (acquisition) and the type 3 hazard (IPO), and the second line is the contrast between the type 2 (liquidation) and the type 3. The high values of the chi-square statistics (equivalently, the low p-value) indicate that neither function is proportional. The proportionality hypothesis must therefore be rejected for each hazard function.

This preliminary analysis already gives some clues to assess the behavior of the hazard and survivor function, and to the behavior that can be expected from some variables. So far, it is established that the survivor functions for each type of exit are distinct, and that the hazard functions, in addition to being clearly different, are not mutually proportional.

Furthermore, it appears that exit by IPO is the least frequent type, followed by exit by liquidation, and finally exit by acquisition, which is the most frequent. This is not surprising considering what was mentioned in the literature review; while IPOs tend to happen exclusively to the most successful firms, acquisition is a more universal exit channel. It is therefore not surprising to see more acquisitions than IPOs.

From the analysis of the sample, it also seems that business angels do have an impact on the exit by reducing the time until an exit occurs. However, the picture is less clear when each exit type is taken separately.

B. Estimation results

The descriptive analysis realized in the previous section gave some interesting indications about the dataset and each exit type. The next step is to analyze the results from the competing risks model, which I will do in this section.

For each type of exit, the model incorporates all the covariates discussed in Chapter 3. The generalized Gamma density function is used as the distribution for the underlying error term. In order to avoid multicolinearity problems[8], the variable SEED, for the financing stage, and the variable WEB, for the industry type, are not included. This means that they in fact become the reference variable, against which the other covariates are to be contrasted.

Besides, I also study the residuals of the models after each estimation. In line with the literature on survival analysis (among others, Kalbfleisch and Prentice, 2002), I evaluate the generalized Cox-Snell residuals (Cox and Snell, 1968), see Appendix 3. Under this method, if the plot of the residuals is close to a straight line with unit slope and zero intercept (i.e. the residuals are exponentially distributed), the distribution is appropriate and the model fits well. I also examine the probability plot, presented in Appendix 4. This topic is discussed in more detail in the next section, but in both cases the model fits well.

The estimates of the model are reported in table 5. To interpret the results, time ratios are also calculated. As mentioned, time ratios are a comparison of rates at which subjects travel the survival curve. The effects of covariates serve to accelerate the passage of time.

[8] Also referred to as the "dummy variable trap", perfect multicolinearity occurs when an exhaustive list of covariates as well as a constant term are present in the regression model. This problem can be avoided by removing either the constant term or one of the dummy variable. The removed dummy then becomes the reference category against which the other categories are assessed. For more details on multicolinearity and the use of dummy variables, refer to: Suits, D. B. (1957). Use of Dummy Variables in Regression Equations. *Journal of the American Statistical Association*. *52*(280), 548–551.

Table 5: Estimation results

Coefficient	Acquisition		Liquidation		IPO	
Accelerated failure time model						
	Estimate	Time Ratio	Estimate	Time Ratio	Estimate	Time Ratio
Early	-0,1657 ***	**18%**	0,3073 ***	**-26%**	-0,5571 ***	**75%**
Expansion	-0,1619 ***	**18%**	0,2632 ***	**-23%**	-0,6518 ***	**92%**
Later	-0,1329 **	**14%**	0,288 ***	**-25%**	-1,0680 ***	**191%**
roundNB	-0,1538 ***	**17%**	-0,0047	**0%**	-0,2316 ***	**26%**
Amount	-0,0001 ***	**0%**	0,0001 ***	**0%**	-0,0002 **	**0%**
Participants	-0,0328 ***	**3%**	0,0446 **	**-4%**	-0,0393 *	**4%**
Angel_dummy	-0,1228 **	**13%**	-0,0382	**4%**	0,297 *	**-26%**
VC_dummy	-0,2316 ***	**26%**	0,0278	**-3%**	0,4632 ***	**-37%**
Milestones_round	-0,0233 *	**2%**	0,0352	**-3%**	-0,0359 ***	**4%**
advertising	0,0762	-7%	0,5327 ***	**-41%**	-0,4486 ***	**57%**
ecommerce	0,4094 ***	**-34%**	0,2068 **	**-19%**	-0,1177	**12%**
social	0,1274 *	**-12%**	0,3705 ***	**-31%**	-0,4983 ***	**65%**
software	0,1956 ***	**-18%**	0,4558 ***	**-37%**	-0,1621	**18%**
hardware	0,5551 ***	**-43%**	0,1908 *	**-17%**	-0,2798 *	**32%**
mobile	-0,0029	**0%**	0,2424 ***	**-22%**	0,1008	**-10%**
enterprise	0,0577	**-6%**	0,684 ***	**-50%**	-0,054	**6%**
network_hosting	-0,1681 *	**18%**	0,2307 *	**-21%**	-0,4433 ***	**56%**
service	0,2622 ***	**-23%**	0,5641 ***	**-43%**	-0,3756 ***	**46%**
education	0,3304 **	**-28%**	12,015 ***	**-100%**	-0,8139 ***	**126%**
games_video	-0,0184	**2%**	0,02	**-2%**	0,1176	**-11%**
IPO_market_global	0,0009	**0%**	-0,0048 ***	**0%**	-0,0026 *	**0%**
IPO_market_tech	-0,0024	**0%**	-0,008 ***	**1%**	-0,0069 *	**1%**

Estimated coefficients for the accelerated failure time model set in the framework of competing risks with three types of exit: acquisition, liquidation, and IPO. The duration is defined as the number of days between the start of a given financing round and the time of the exit. In the case of non-exited rounds, their duration is considered as right-censored at the date of the 31st December 2013 and included in the model. The specified density distribution of the model is the generalized Gamma density distribution. A ***, ** and * indicates that the coefficient is significant at a 1%, 5% and 10% level respectively.

1. Acquisition exit

The results show that the "Early" group will experience an acquisition 18% faster than the "Seed" group (which serves as the reference group in this case). In other words, belonging to the "Early" group will accelerate the time until the acquisition by 18%. The percentage is the same for “Expansion” group, and it becomes 14% for “Later” group. All three covariates are statistically significant with *p*-values well beyond the 5% threshold.

This result is no surprise: projects that are less advanced take more time to be acquired than those that are more mature. On the other hand, this reduction in the time to acquisition is less pronounced in the case of companies in the later stage. A possible reason for this may be that the very successful projects do not exit by acquisition but by IPO, a preferred route for entrepreneurs and venture capitalists. Another possibility could be that, after some time, when the company is more mature, it starts showing signs of weakness, signs that were not apparent when the project was still budding. Overall however, the stage variables behave as expected from the literature review.

The number of rounds is also a statistically significant variable, with a strong effect on the survival time. For each additional financing round, the time to exit of the firm is accelerated by 17%. This makes sense: if a firm consistently attracts interest from investors round after round, it must mean that there are strong hopes that it will become successful, otherwise investors would just stop funding it.

Because more rounds also mean more money and, usually, more investors, the value added by each factor pushes the company towards a favorable exit. This is confirmed by the PARTICIPANTS and AMOUNT variables, both of which are significant. Each additional investor in the funding round decreases the time to acquisition by 4%. This supports the rationales for deal syndication, whether because a syndicate of investors selects projects better, or because company benefits from an increased pool of financial or other resources.

The coefficient for the AMOUNT variable appears to be 0; this is be due to the fact that because the variable is expressed in $1000, adding a single thousand dollar makes no measurable difference and does not impact the timing of the exit in a meaningful way. However, the sign of the coefficient indicates that the amount has a positive impact. Thus, a larger amount leads to a decreased time to exit. This is in line with previous studies.

As anticipated from the descriptive analysis, the presence of business angels and the presence of several venture capitalists has a significant impact on the acquisition exit, by accelerating the survival time by 13% and 26% respectively. It thus appears that when business angels are present, firms tend to undergo an acquisition faster than when they are not present. Moreover, since angels tend to be present at earlier stages, it can be concluded that their presence has a long-lasting impact on the firm, i.e. from the point they start being involved until the company is acquired.

However, the effect for variable VC_DUMMY is much stronger. Whether because of the better selection or the greater added value, the results show that when more than one VC is involved, the firm tends to exit 26% faster. This is confirmed when plotting the CIF for the group where two or more VCs are present against the other group.

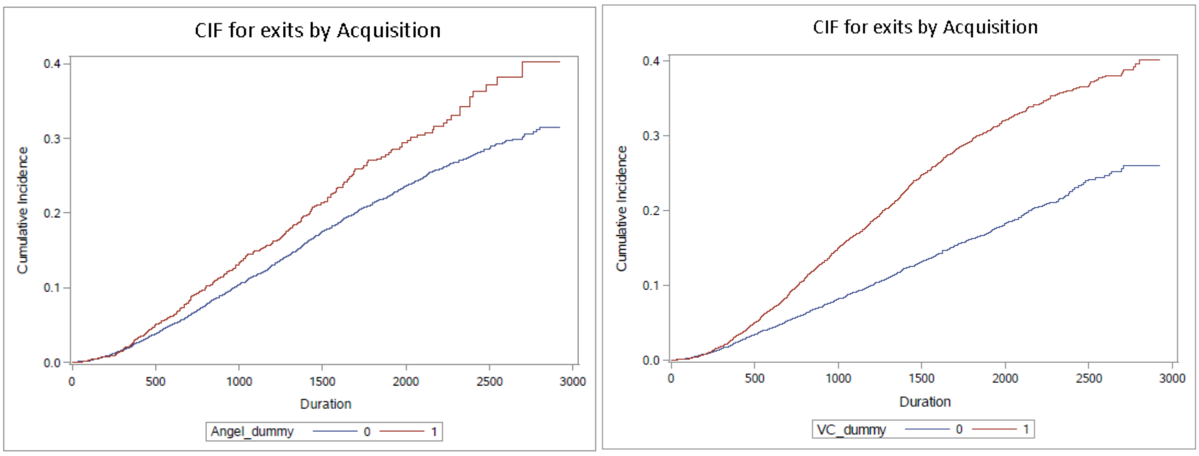


Fig 13: CIF plots for exit by acquisition

As depicted by Fig 13, the CIF when the ANGEL_DUMMY and VC_DUMMY variables are equal to 1 is far above the CIF when they are equal to 0. This is a clear sign that business angels and consortiums of venture capitalist increase the occurrence of exits by acquisition.

Simulations show, that for the first financing round of a company, when only angels are involved, increasing the number of participants from two to four reduces the exit time by 6%. When the number of VCs involved is increased from two to four, the time to exit is reduced by 8%. This confirms that an increased number of participants do reduce the time to exit, and that this effect is stronger when the participants are VCs. Moreover, when there are only two participants that are VCs, firms exit 9% faster than when the two investors are angels; this ratio becomes 11% when the number of VCs in increased to four.

Estimation results also show that there is a non-negligible industry effect. The negative time ratios indicate that, by contrast to web companies, companies in other industries tend to exit more slowly (with the exception of companies specialized in network hosting, although the effect here is only marginally significant). The most extreme difference is between network hosting and hardware companies. Their relative time ratio is more than 2 (computed as $e^{0.5551}/e^{-0.1681}$), meaning that network hosting companies are acquired two times faster than hardware companies.

The number of milestones achieved in the previous rounds is only marginally significant, with an estimate close to 0 and a time ratio of 2%, i.e. for each milestone achieved in the previous rounds the firm's time to exit is reduced by 2%. This is in accordance with what was expected from the literature review.

The results for both IPOs variables are surprising, however. Indeed, according to the literature review these variables should have some sort of impact, for example, that a very active IPO market might encourage more companies to go public instead of being acquired, but this is not confirmed by the model. The "hotness" of the IPO market makes no difference to the time needed for companies to be acquired, i.e. the state of the IPO market neither facilitates nor hinders an exit by trade sale.

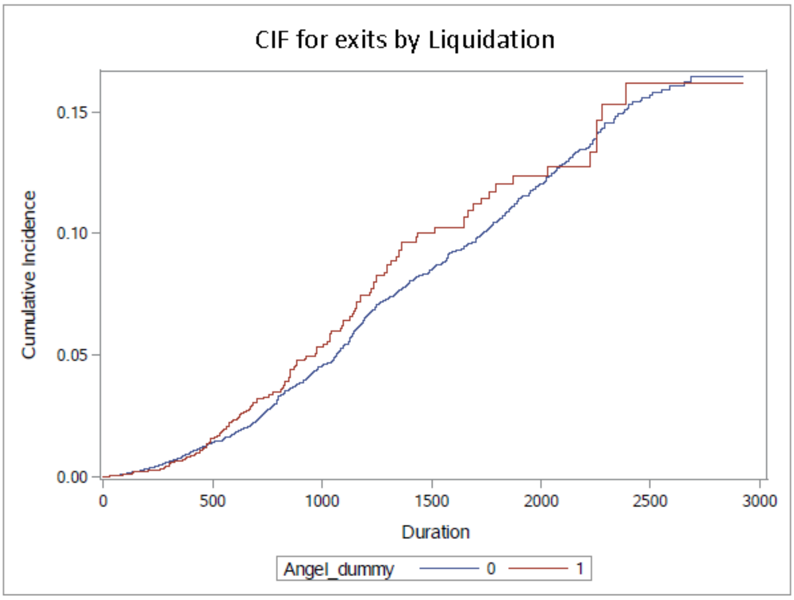
Overall, the results for exit by trade sale are quite consistent and in accordance with what was expected from the literature review and similar studies.

2. Liquidation exit

Once again, the stage variables are significant. As expected, the very early-stage companies are the most vulnerable to liquidation. Reaching the "early" stage delays any liquidation by 26%, a ratio that stays more or less uniform for the other financing stages.

A number of variables that were significant when estimating the results for exit by acquisition are no longer significant. The number of rounds is one such variable. This may indicate that, for example, if the company's business model is flawed or ineffective, it does not matter how many financing rounds the company undergoes: they will not prevent liquidation. The number of milestones is another variable that is no longer significant. One possible reason might be that achieving technological progress also gives information about the future product. Even if the company reaches milestones, if its product or business model does not offer opportunities for profitability and viability the company will be at risk of liquidation.

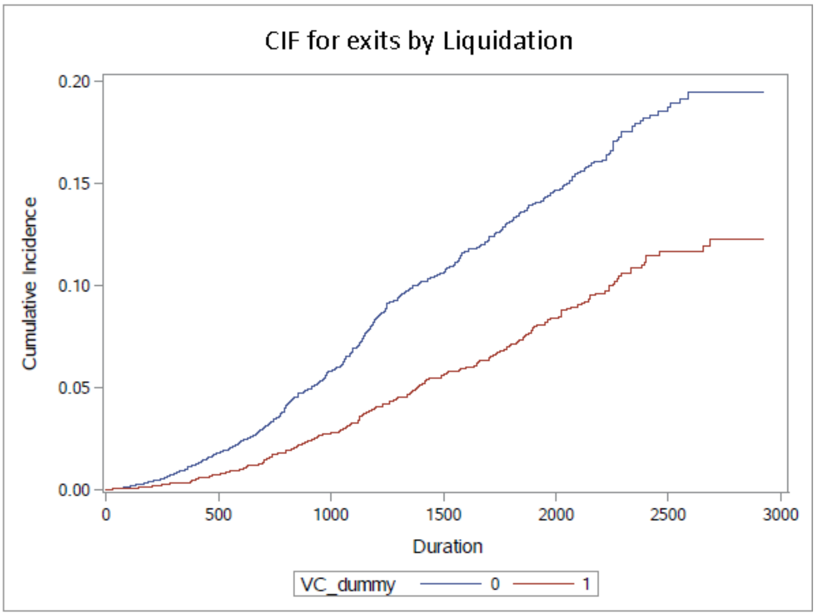
As hinted by the descriptive analysis, the ANGEL_DUMMY and VC_DUMMY are not significant in the case of exit by liquidation. To confirm these results, I plotted the CIF for the two variables and estimated the Gray test for equality of cumulative incidence functions. As depicted in Fig 14, the plot indicates that there is no discernible difference between the groups for the two levels of the ANGEL_DUMMY variable, which is corroborated by the hypothesis test.



Chi-Square	DF	Pr > Chi-Square
1.28430	1	0.2571

Fig 14: CIF plot for the variable ANGEL_DUMMY for exit by liquidation

However, the results are very different for venture capitalists as depicted by Fig 15.



Chi-Square	DF	Pr > Chi-Square
84.7555	1	<.0001

Fig 15: CIF plot for the variable VC_DUMMY for exit by liquidation

The plot clearly shows two distinct cumulative incidence functions. The function corresponding to investment rounds where two or more VCs are present lies far below the other. Furthermore, the test rejects the hypothesis of equality between the two CIFs. This indicates that the frequency of liquidation when at least two VCs are present is greatly reduced.

One possible reason for such difference between the two variables may be due to the information asymmetry problem. As mentioned in the literature review, since business angels invest in very early-stage companies, when the project still has to prove its viability, the risk of failure is higher. Conversely, VCs tend to invest later, so they often have more information about the company. The results indicating that business angels have no impact on liquidations, while venture capitalists have a substantial effect, may simply come from the fact that since they have more information, VCs can better select the successful project – or at least those that will not liquidate. Note that the angel variable only denote the group of investment rounds in which business angels are present, against the group where only VCs are present. It would not be surprising to see that a fair number of companies that liquidated never made it to the point where they would be of interest for VCs in the first place.

The amount and number of participants are two significant variables. Both of them have a positive impact by delaying a potential liquidation. While the effect is not measurable for the amount, for the number of participants, however, it corresponds to a 4% duration increase for each additional participant.

The variables for the IPOs are highly significant, however the estimates are very close to 0, meaning that their effect on the time to exit are tenuous. The effect is more pronounced for the high-tech IPO market. The negative signs indicate that when the number of IPOs increases, the time to exit is reduced, i.e. as the IPO markets become more active, companies liquidate faster. It may seem surprising but in fact it has been shown that in "bubble" periods companies tend to raise more money than in depressed periods (Block, J., Sandner, 2009). Since the sample perfectly covers the period before and after the height of the 2008 crisis, this is another example of this phenomenon. When markets are optimistic, venture capitalists tend to give more money to firms that otherwise would not receive as much. And because they do not perform as initially expected they also tend to liquidate much faster than they would if markets were more circumspect. Thus leading to positive time ratios, indicating accelerated liquidations when the number of IPOs increases.

The effect of the industry type variable varies greatly. What is clear given the fact that all the time ratios are negative is that the web industry tends to liquidate much faster than any other industry, education companies are, however, the slowest to liquidate.

Globally, the estimation results for exit by liquidation behave as expected from the literature review and the descriptive analysis. Whereas most covariates accelerated the life of observation in the case of exit by acquisition, in this case they tend to make it slower, i.e. they tend to delay the potentiality of liquidation. Interestingly, while the AFT model determined that the venture capitalist dummy variable is not significant, the CIF shows that venture capitalists have a strong effect in reducing the frequency of liquidation.

3. IPO exit

The results for exit by IPO are to be treated with caution given the rarity of this exit in the dataset. The sample size for IPO exits is much smaller than for the other two exit types.

The estimation results indicate that the stage of the company has an extremely strong effect on its time to exit. Compared to companies in a round occurring at the "seed" stage, those in a round occurring at the "early" stage exit 75% faster, those in a round occurring at the "expansion" stage exit 92% faster, and finally those in a round occurring at the "later" stage exit 191% faster. This confirms that mature companies with a more sophisticated project are closer to an IPO than companies that are in a more experimental phase.

The milestone variable also points out that more advanced companies exit faster: an acceleration of 4% for each milestone achieved. Likewise the amount invested increases.

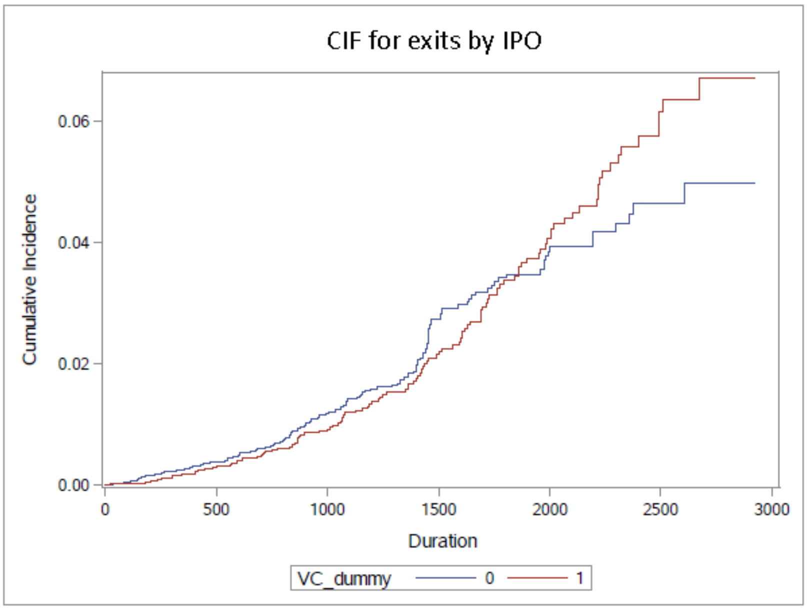
Similarly, the number of rounds, which also gives a rough indication of the stage of the company, greatly accelerates the occurrence of an IPO, by 26% for each additional financing round. This ratio was only 17% for acquisitions.

The number of participants decreases the time needed for a company to go public, although this variable is only marginally significant.

There are wide variations across industries. The education and social industries exhibit a tendency to exit faster than web companies. By contrast with exit by acquisition, the majority of time ratios are positive, indicating that the web industry takes more time to exit by IPO than other industries.

This time the IPO variables are significant, although only at a marginal level, and the sign of the estimates suggest that an active IPO market reduces the time until a company goes public.

Regarding the type of investor, the results are quite surprising. The presence of two or more venture capitalists appears to be highly significant, delaying a potential IPO by 37% (compared to when only one or no VC are present). However, the plot the CIF of this variable does not confirm the results, as seen on Fig 16.

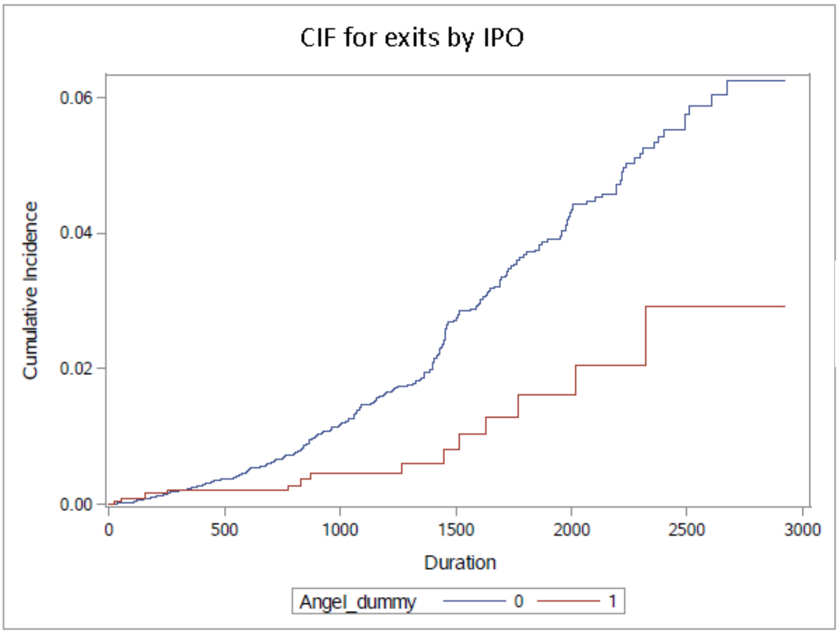


Chi-Square	DF	Pr > Chi-Square
0.025781	1	0.8724

Fig 16: CIF for VC_DUMMY in the case of IPO exit

The results found by this method are not conclusive. The Gray test of equality fails to reject the hypothesis of equality among the two groups, i.e. it has to be concluded that there is not enough evidence from the dataset to reject the possibility that the two functions are the same, using this method at least.

However, the CIF for business angels, Fig 17, shows a clear difference between the group where at least one angel is present and the other group, and the test also rejects the hypothesis of equality with a strong level of confidence. This confirms the results from the model that indicated that the presence of business angels increased the time to an IPO by 26%.



Chi-Square	DF	Pr > Chi-Square
11.4551	1	0.0007

Fig 17: CIF for ANGEL_DUMMY in the case of IPO exit

The rough and discontinuous appearance of the curves is due to the fact that IPOs are a less frequent type of exit, and especially those in which at least one business angel is present. The underrepresentation of IPOs in the dataset might indicate that the results may not be reliable and does not give the possibility to draw meaningful conclusions for these variables.

C. Goodness of fit, robustness and heterogeneity

This section will cover the methods that were used to assess the goodness of fit of the model as well as the robustness of the results, and ends by giving some more details on the heterogeneity that is expected in this study.

The first method used to assess the goodness of fit is based on the Cox-Snell residuals. According to Cox and Snell (1968) if the residuals are exponentially distributed it is a sign that the data are well fitted by the model. To test this hypothesis, the cumulative hazard function of the residuals can be plotted against a benchmark line with slope equal to 1 and intercept of 0. Fig 18 plots these functions for exit by acquisition. The graphs indicate that the cumulative hazard function is very close to the benchmark line. The more volatile segments on the right-hand side are caused by some outliers in the residuals, however, hundreds of residuals are aligned almost perfectly on the line. The results for liquidation and IPO exit are reported in Appendix 3 and are quite similar.

Globally, the Cox-Snell analysis of residuals tends to indicate that the model fits the data closely.

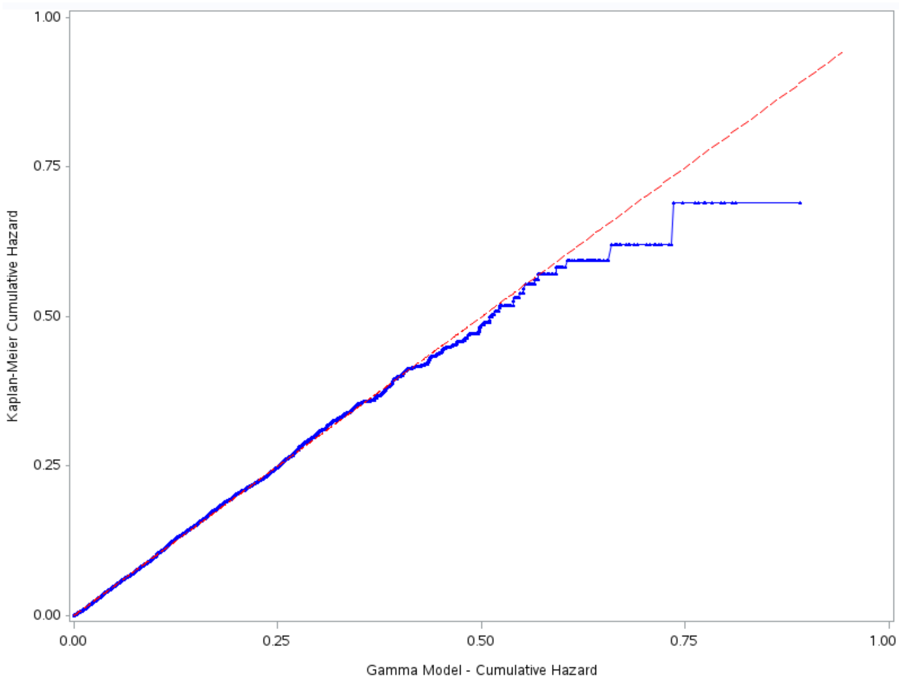


Fig 18: Cox-Snell residuals for exit by acquisition

The other method of analysis of the goodness of fit is the probability plot, which applies a transformation to the survivor function estimated using a non-parametric Kaplan-Meier method adjusted to take into account the effect of the covariates. If the specified model is correct, the data on the plot should form a straight line.

As shown by Fig 19, the probability plot for the acquisition exhibits a linear pattern, and all the non-parametric estimates lie within the 95% confidence interval. Probability plots for the other exit types are provided in Appendix 4.

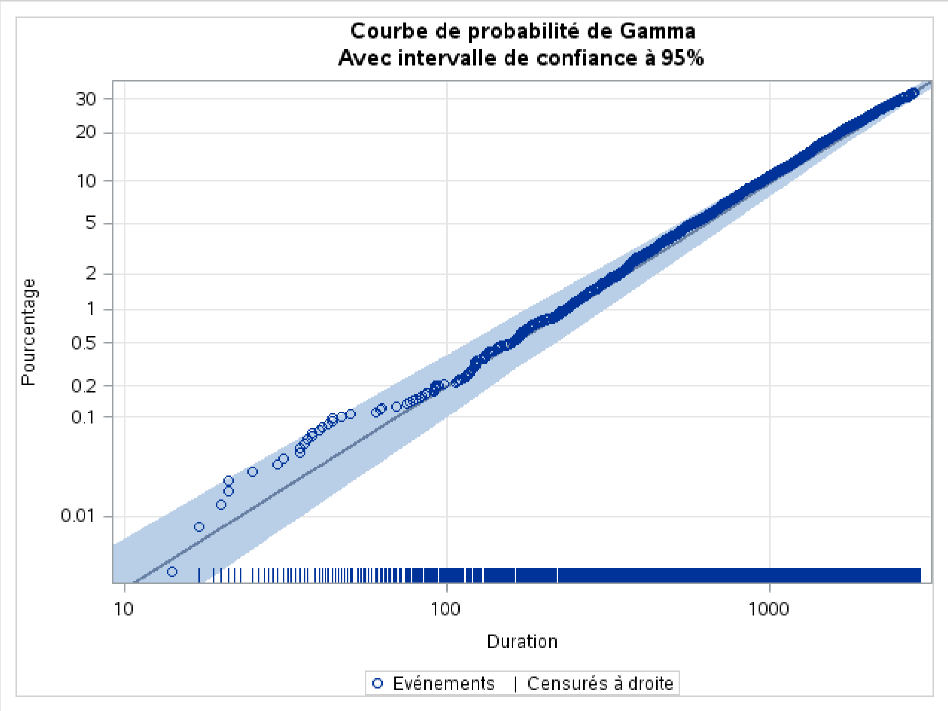


Fig 19: Probability plot for exit by acquisition

Overall, the probability plots show that the model fits well and corroborate the choice of the Gamma distribution.

In order to check the robustness of the model, I also applied a Cox proportional hazard model. The results of this model are reported in Appendix 5.

While the estimates of the Cox model are more extreme than the estimates of the AFT model, they do reflect the same trends. Positive time ratios in the AFT model always translate into positive hazard ratios in the Cox model. Furthermore, the two models also display some degree of proportionality. For example, for exit by acquisition the time ratio of the e-commerce industry is higher than that of the software industry, and this is reflected in the hazard ratios of the Cox model.

Note, however, that the results of the two models have different interpretations. The hazard ratio is the ratio of the estimated hazard for one group against the estimated hazard for the other group. For example, in the case of acquisitions, the hazard ratio between the “seed” group and “early” group is 28%, meaning that the likelihood that companies in the "early" group experience an acquisition is 28% higher than the likelihood of companies in "seed" group (the reference group in this case).

Before concluding, let’s consider the impact that unobserved heterogeneity may have on the estimation results. As previously mentioned, it would be impossible to build a dataset that captures all the information about the investments, the companies and their environment. A certain level of unobserved heterogeneity is thus to be expected.

In the presence of unobserved heterogeneity the estimated hazard functions tend to decline with time, even if the actual hazard is not reduced for any observation (Heckman and Singer, 1985). According to Allison (2010), when an increasing hazard function is found, this can be interpreted as a sign that the real hazard increases over some interval of time for a portion of the observation. In the previous section, the hazard functions for the dataset were plotted using the kernel smoothed method and indicated that the hazard for the acquisition exit showed an upward trend at the end of the time interval. The hazard function for IPO remained stable and the one for liquidation decreased at the end of the period.

The case of decreasing hazard functions is ambiguous; this is why frailty models have been developed. Such models try to separate the hazard function from unobserved heterogeneity by adding a random disturbance term into a Weibull hazard model. However, such models are extremely sensitive to the choice of a distribution for the error term (Allison, 2010).

Gail et al. (1984) found that unobserved heterogeneity attenuates the regression coefficients to zero. However, standard errors and test statistics are not biased. Therefore, it is still possible to test the hypothesis that a coefficient is null, even when unobserved heterogeneity is present. I have conducted such tests on the regression coefficient and found that this hypothesis should be rejected.

Furthermore, studies have shown that when unobserved heterogeneity is present, the use of survival analysis models result in regression coefficients as good as models that include a special parameter dedicated to capturing such unobserved heterogeneity (Liu, 2014).

This is one reason why no frailty variable was included in the model.

VII. Conclusion and future directions

Through the use of survival analysis methods set in the framework of competing risks, I was able to analyze simultaneously the two components that constitute the exit strategy of venture capitalist and business angels: the time to exit and the type of exit.

Meaningful conclusions that can be reached after examining the survivor and the hazard functions of the financing rounds in the dataset. First, it appears that trade sales are the most frequent exit route. The likelihood of a trade sale starts rising as soon as the financing round begins, it reaches levels higher than any other type of exit, with a peak at 1.500 days. The likelihood of liquidations is not as high, however it starts increasing soon after the beginning of the round to reach a plateau at 1.000 days until 2.500 days, it starts to decline sharply afterwards. The least frequent exit route is the IPO, the hazard function for IPO reaches its maximum at 1.500 days and levels off afterwards.

The results of the model also show that a series of variable has been consistently significant.

First, the parameters that relate to the stage of development play an important role in the time an investment needs to be exited. The more advanced the company, the closer it is from a favorable exit. It is the opposite when considering exits by liquidation, in which case the advancement of the company actually delays the occurrence of a potential liquidation. This therefore confirms that investments in very early-stage projects are riskier.

Second, the syndication of venture capital deals has a positive impact on the duration of investment. When more investors are involved, the investment time tends to be shorter (except when considering liquidation in which case the life of the company is extended). This indicates that the advantages of syndication are real, but it does not identify the exact cause of such benefits. More precisely, it does not give information on whether it is the better selection of projects by the syndicate or the increased pool of management skills that helps reduce the time to exit.

Next, it appears that web-related companies exhibit some particular tendencies. For example, they are quicker to liquidate than any other internet industry, but they are also the ones that tend to be acquired faster as well. Results are more diverse when examining the time to IPO, in which case other industries tend to do better.

Other variables such as the amount invested and the development milestones of the companies can positively impact the duration of investments.

This answers the first two research questions about the significance of the variables and their impact. My results also shed some light on the interactions between the types of investors and the exit strategy.

The presence of at least one business angel makes a trade sale much more likely and reduces the time for the company to get acquired by 13%. In the cases of exits by liquidation or IPO however the impact of business angels is not clear since the AFT model and the CIF functions give contradictory results. It seems however that their presence among the investors does not make a difference that is statically significant. One reason for that is that, since business angels invest earlier, when the project still has to prove its viability, the risk of failure is higher. Conversely, when venture capitalists invest they often have more information about the company, they have therefore the opportunity to pick projects that have already shown signs of viability. The results indicating that business angels have no impact on liquidations, while venture capitalists have a substantial effect, may simply come from the fact that since they have more information, venture capitalists can better select the successful projects.

When more than two venture capitalists are involved in the deal, it tends to make trade sale both more likely and faster. Their presence greatly reduces the likelihood of liquidation and delays its occurrence by 26%.

This gives leads to answer the last research question; it is clear that business angels have an impact on the type and timing of the exit. But generally the impact generated by presence of at least two venture capitalists outweighs the impact of business angels. Moreover, their impact is positive since they increase the likelihood of a favorable exit while reducing the time until such exit occurs.

However, results on the type of investor regarding the IPO exit are contradictory and conclusions might not be meaningful. The reason for this may stem from the fact that IPOs are underrepresented in the dataset. Since the results were consistent for the other types of exit, future research on IPO exits should consider the distinction between angel investors and venture capitalists when investigating the rationales behind this exit.

Regarding the sample, it would be very interesting to see if the results of this study are confirmed when using a different database. Although CrunchBase does not benefit from the scrutiny of other databases specialized on venture capital investment, such as Thomson Reuters's VentureXpert or Venture Source from Dow Jones, it still represents a promising alternative (Kaplan and Lerner, 2015). Another interesting development would be to apply the same methodology to the fully up-to-date information from CrunchBase. Indeed, my dataset stops at the end of 2013, but there are three and a half year of additional data available from this website, unfortunately those are not free.

Furthermore, the results seem to indicate that IPO markets only have marginal impact on the exits of internet companies. Besides, many studies have examined the consequences of the internet bubble on the venture capital industry, but comparatively, only few have investigated the impact of the more recent crisis of 2008. In 2013, the number of IPOs in the US reached the level it had in 2007, but polls frequently find that people are pessimistic on the future prospects of the economy[9]. A study that uses survival analysis methods to examine the impact of the crisis on the time to exit of venture capital investments could lead to some interesting results, and even more so if focusing on the internet industry. A study using the data collected by the Global Entrepreneurship Monitor[10] on the state of entrepreneurship around the world could also be considered in order to examine the effects of the 2008 period on entrepreneurial ventures or the desire for entrepreneurship.

[9] See for example :
http://www.theatlantic.com/politics/archive/2016/02/america-economic-outlook-heartland/458727/
http://www.wsj.com/articles/wsj-nbc-poll-finds-widespread-economic-anxiety-1407277801

[10] http://www.gemconsortium.org/

VIII. References

Allison, P. D. (2010). *Survival analysis using SAS: a practical guide*, second edition. Sas Institute.

Bascha, A. & Walz, U. (2001). Convertible securities and optimal exit decisions in venture capital finance. *Journal Of Corporate Finance*, *7*(3), 285-306.

Bayar, O. (2006). IPOs versus Acquisitions and the IPO Valuation Premium Puzzle: An Empirical Analysis. *Unpublished Working Paper*, *Boston College.*

Bergemann, D., & Hege, U. (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking and Finance*, *22*, 703-35.

Brau, J., Francis, F., & Kohers, N. 2003. The Choice of IPO versus Takeover: Empirical Evidence. *Journal of Business*, *76*, 583-612.

Black, B., & Gilson, R. (1998). Venture capital and the structure of capital markets: banks versus stock markets. *Journal of Financial Economics*, *47*, 243–277.

Black, B., & Gilson, R. (1999). Does venture capital require and active stock market?. *Journal of Applied Corporate Finance*, *11*, 36–48.

Block, J., & Sandner, P. (2009). What is the effect of the financial crisis on venture capital financing? Empirical evidence from US Internet start-ups. *Venture Capital*, *11*(4), 295-309.

Bloomfield, S. (2008). *Venture capital funding*. London: Kogan Page.

Brander, J., Amit, R., & Antweiler, W. (2002). Venture capital syndication: improved venture selection vs the value-added hypothesis. *Journal of Economics and Management Strategy*, *11*, 423–452.

Brau, J., Francis, B., & Kohers, N. (2003). The Choice of IPO versus Takeover: Empirical Evidence. *Journal of Business*, *76*(4), 583-612.

Bygrave, W. & Timmons, J. (1992). *Venture capital at the crossroads*. Boston, Mass.: Harvard Business School Press.

Bygrave W. D., Hay M., Peeters J. (1999). *The venture capital handbook*. Harlow [England]: Financial Times Prentice Hall.

Cantor, A. (2003). *Sas Survival Analysis Techniques for Medical Research*. SAS Institute, 111–150.

Cleves, M., Gould, W.W., & Gutierrez, R. (2004). *An introduction to survival analysis using Stata*. College Station, Tex: Stata Press.

Cochrane J. (2001). The Risk and Return of Venture Capital. *SSRN Electronic Journal*. Downloadable at: http://dx.doi.org/10.2139/ssrn.253798.

Cox, D. R. (1972). Regression Models and Life Tables" (with discussion). *Journal of the Royal Statistical Society*, *34*, 187-220.

Cox, D.R., & Snell, E.J. (1968). A general definition of residuals (with discussion). *Journal of the Royal Statistical Society*, *30*, 248–275.

Cox, D. R,. & Oakes, D. (1984). *Analysis of Survival Data*. London: Chapman & Hall.

Cumming, D. (2007). Government policy towards entrepreneurial finance: Innovation investment funds. *Journal of Business Venturing*, *22*, 193-235.

Cumming, D. (2008). Contracts and exits in venture capital finance. *Review of Financial Studies*, *21*(5), 1947-1982.

Cumming, D., & MacIntosh, J., (2001). Venture capital investment duration in Canada and the United States. *Journal of Multinational Financial Management*, *11*, 445–463.

Cumming, D., & MacIntosh J., (2003a). A Cross-Country Comparison of Full and Partial Venture Capital Exits. *Journal of Banking and Finance*, *27*(3), 511-548.

Cumming, D., MacIntosh J., (2003b). Venture capital exits in Canada and the United States. *University of Toronto Law Journal*, *53*, 101–200.

Das, S. R., Jagannathan, M., & Sarin, A. (2003). Private Equity Returns: An Empirical Examination of the Exit of Venture-Backed Companies (Digest Summary). *Journal of Investment Management*, *1*, 1152-177.

Denis, D.J., (2004). Entrepreneurial finance: an overview of the issues and evidence. *Journal of Corporate Finance, 10,* 301–326.

Félix, E., Pires, C., & Gulamhussen, M. (2012). The exit decision in the European venture capital market. *Quantitative Finance*, *14*(6), 1115-1130.

Fleming, G. (2002). Venture capital returns in Australia. *Venture Capital*, *6*(1), 23-45.

Gail, M. H., Wieand, S., & Piantadosi, S. (1984). Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika, 71*(3), 431-444.

Giot, P., & Schwienbacher, A. (2005). IPOs, trade sales and liquidations: modelling venture capital exits using survival analysis. *CORE Discussion Papers*; 2005/13.

Gompers, P. A. (1995). Optimal Investment, Monitoring, and the Staging of Venture Capital. *The Journal of Finance, 50* (5), 1461–1489.

Gompers, P. A. (1996). Grandstanding in the venture capital industry. *Journal of Financial Economics, 42*(1), 133–156.

Gompers, P., Lerner, J. (1999a). What drives venture capital fundraising?. *SSRN Electronic Journal*. NBER working paper 6906. http://dx.doi.org/10.2139/ssrn.57935

Gompers, P. A., & Lerner, J. (1999b). *The Venture Capital Cycle*. Cambridge, Mass.: MIT Press.

Gray, R. (1988), A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. *The Annals of Statistics, 16*, 1141–1154.

Heckman, J. J. and Singer, B. (1985), Social Science Duration Analysis. In *Longitudinal Studies of Labor Market Data*, ed. J. J. Heckman and B. Singer, New York: Cambridge University Press.

Hochberg, Y., Ljungqvist, A., & Lu, Y. (2007). Whom You Know Matters: Venture Capital Networks and Investment Performance. *The Journal Of Finance*, *62*(1), 251-301.

Jenkins, S.P. (2005). *Survival analysis*. Unpublished manuscript, Institute for Social Science and Economic Research, University of Essex, Colchester, UK. Downloadable at: https://www.iser.essex.ac.uk/files/teaching/stephenj/ec968/pdfs/ec968lnotesv6.pdf

Kalbfleisch, J. D. & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data,* Second Edition, New York: John Wiley & Sons, Inc.

Kaplan, S. N., & Lerner, J. (2015). Venture capital data: Opportunities and challenges. In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*. University of Chicago Press.

Kerr, W. R., Lerner, J., & Schoar, A. (2010). *The Consequences of Entrepreneurial Finance: A Regression Discontinuity Analysis*. Unpublished manuscript. Harvard Business School, Entrepreneurial Management Working Paper No. 10-086, Harvard.

Kaplan, S. N., & Lerner, J. (2015). Venture capital data: Opportunities and challenges. In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*. University of Chicago Press.

Lau, B., Cole, S. R., & Gange, S. J. (2009). Competing risk regression models for epidemiologic data. *American journal of epidemiology*, *170*(2), 244-256.

Lee, E. T., & Wang, J. (2003). *Statistical methods for survival data analysis* (Vol. 476). John Wiley & Sons.

Lerner, J. (1994a). Venture capitalists and the decision to go public. *Journal of Financial Economics*, *35*, 293–316.

Lerner, J. (1994b). The Syndication of Venture Capital Investments. Financial Management, 23(3), 16-27.

Lerner, J. (1999). The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program. *Journal of Business*, *72*(3), 285-318.

Lockett, A., Murray G., & Wright M. (2002). Do UK venture capitalists still have a bias against investment in new technology firms?. *Research Policy*, *31*, 1009-1030.

Lin, D. Y., Wei, L. J., & Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika*, *80*(3), 557-572.

Liu, X. (2014). Survival Models on Unobserved Heterogeneity and their Applications in Analyzing Large-scale Survey Data. *Journal of Biometrics & Biostatistics*, *5*, 191. http://doi.org/10.4172/2155-6180.1000191

Nahata, R. (2008). Venture capital reputation and investment performance. *Journal Of Financial Economics*, *90*(2), 127-151.

Madill, J. J., Haines, Jr, G. H., & Riding, A. L. (2005). The role of angels in technology SMEs: A link to venture capital. *Venture Capital*, *7*(2), 107-129.

Michelacci, C., & Suarez, J. (2004). Business creation and the stock market. *The Review of Economic Studies*, *71*(2), 459-481.

Murray, G., & Lott, J. (1995). Have venture capitalists a bias against investment in new technology firms?. *Research Policy*, *24*, 283-299.

Murray, G. (1998). A policy response to regional disparities in the supply of risk capital to new technology-based firms in the European Union: The European Seed Capital Fund Scheme. *Regional Studies*, *32*(5), 405-420

Ozmel, U., Robinson, D., & Stuart, T. E. (2013). Strategic alliances, venture capital, and exit decisions in early stage high-tech firms. *Journal of Financial Economics*, *107*(3), 655-670.

Paul, S., Whittam, G., & Wyper, J. (2007). Towards a model of the business angel investment process. *Venture Capital*, *9*(2), 107-125.

Pintilie, M. (2006). *Competing Risks: A Practical Perspective*, New York: John Wiley & Sons, Inc.

Poulsen, A. B., & Stegemoller, M. (2008). Moving from private to public ownership: selling out to public firms versus initial public offerings. *Financial Management*, *37*(1), 81-101.

Ruhnka, J. C., Young, J. E. (1987). A venture capital model of the development process for new ventures. *Journal of Business Venturing*, *2*(2), 167-184.

Sapienza, H. J., Manigart, S., & Vermeir, W. (1996). Venture capitalist governance and value added in four countries. *Journal of Business Venturing*, *11*(6), 439-469.

Singh, R., & Mukhopadhyay, K. (2011). Survival analysis in clinical trials: Basics and must know areas. *Perspectives in Clinical Research*, *2*(4), 145–148.

Sohl, J. (2011). The angel investment market in 2010: a market on the rebound. Center for Venture Research, University of New Hampshire.

Schwienbacher, A. (2002). An empirical analysis of venture capital exits in Europe and in the United States, Mimeo, University of Amsterdam.

Schwienbacher A. (2008a). Venture Capital Investment Practices in Europe and in the United States. *Financial Markets and Portfolio Management*. *22*(3), 195-217.

Schwienbacher A. (2008b). Innovation and Venture Capital Exits. *Economic Journal*, *118*(533), 1888-1916.

Schwienbacher A. (2009). Venture Capital Exits. In *Companion to Venture Capital*, New York: John Wiley and Sons, Inc.

Wong, A., Bhatia, M., & Freeman, Z. (2009). Angel finance: the other venture capital. *Strategic Change*,*18*, 221–230.

Vos, E., Yeh, A. J. Y., Carter, S., & Tagg, S. (2007). The happy story of small business financing. *Journal of Banking & Finance*, *31*(9), 2648-2672.

Zhang, D. (2006). *Analysis of Survival Data*. Unpublished manuscript. Department of Statistics, North Carolina State University, NC. Downloadable at: http://www4.stat.ncsu.edu/~dzhang2/st745/chap5.pdf

IX. Appendix

1. Appendix 1: Test of the proportional hazard assumption

A Pr > MaxAbsVal coefficient lower than 0,10 indicates a rejection of the proportional hazard assumption for the tested covariate with a 90% confidence level.

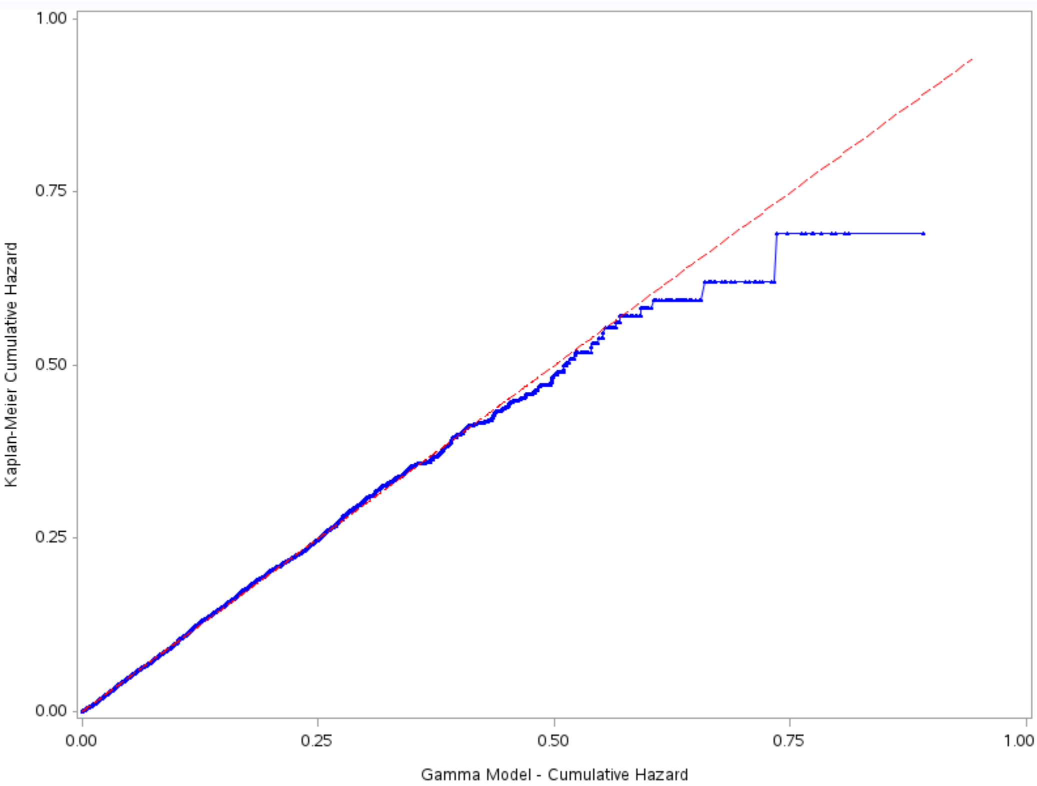
Test de la borne supérieure pour l'hypothèse des risques proportionnels				
Variable	**Valeur absolue maximum**	**Répétitions**	**Souche**	**Pr > MaxAbsVal**
Early	0.9296	1000	1690952018	0.5210
Expansion	1.7028	1000	1690952018	0.0790
Later	0.6953	1000	1690952018	0.8950
roundNB	1.3695	1000	1690952018	0.1040
Amount	1.4255	1000	1690952018	0.0560
Participants	1.7713	1000	1690952018	0.1090
Angel_dummy	0.9735	1000	1690952018	0.4980
VC_dummy	2.5515	1000	1690952018	0.0030
Milestones_round	0.6556	1000	1690952018	0.6700
advertising	1.1428	1000	1690952018	0.2040
ecommerce	0.6187	1000	1690952018	0.8160
social	1.6391	1000	1690952018	0.0270
software	0.7479	1000	1690952018	0.8120
hardware	1.1891	1000	1690952018	0.1280
mobile	1.1268	1000	1690952018	0.2260
enterprise	0.8101	1000	1690952018	0.6410
network_hosting	1.1729	1000	1690952018	0.1560
service	0.8856	1000	1690952018	0.4720
education	1.0964	1000	1690952018	0.1550
games_video	1.7685	1000	1690952018	0.0120
IPO_market_global	7.1584	1000	1690952018	<.0001
IPO_market_tech	6.6239	1000	1690952018	<.0001

2. Appendix 2: Test of equality the of the survival functions

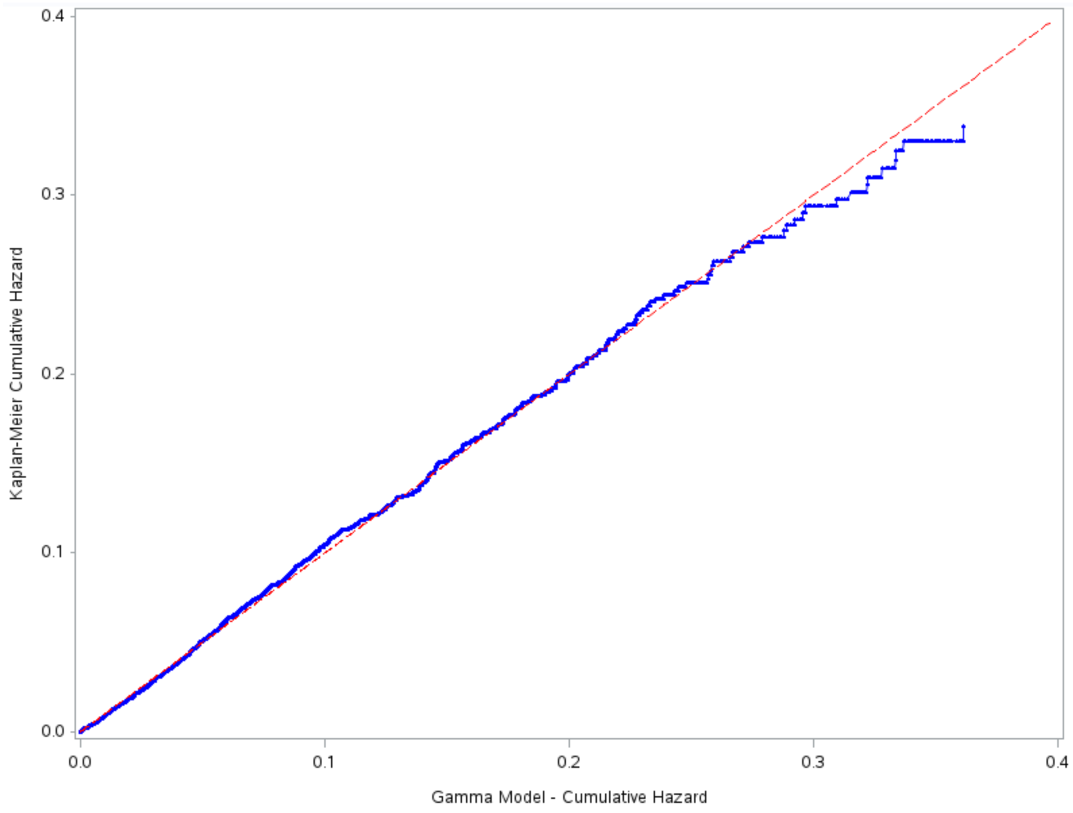
Test d'égalité sur niveaux de discrétisation			
Test	Khi-2	DDL	Pr > khi-2
Log-rang	1684.5930	2	<.0001
Wilcoxon	1478.8219	2	<.0001
Tarone	1650.3849	2	<.0001
Peto	1697.8310	2	<.0001
Peto modifié	1697.8338	2	<.0001
Fleming(1)	1697.8178	2	<.0001

3. Appendix 3: Cox-Snell residuals for each type of exit

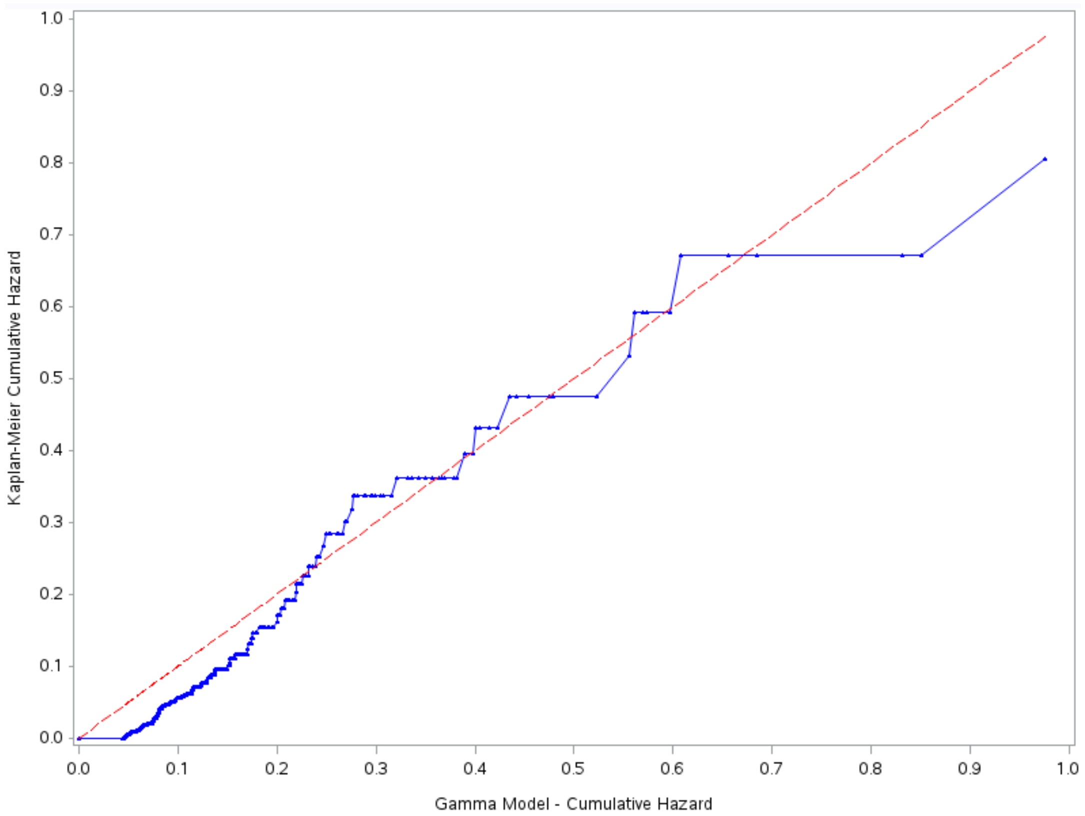
- Cox-Snell residuals for acquisition exit:



- Cox-Snell residuals for liquidation exit:

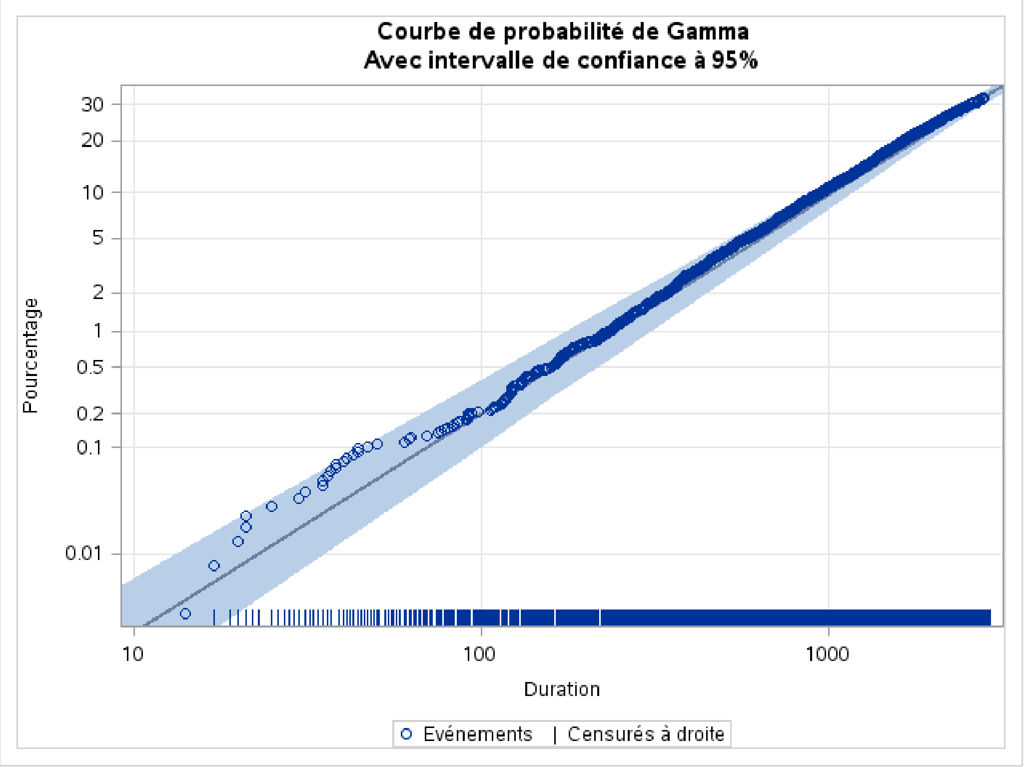


- Cox-Snell residuals for IPO exit:

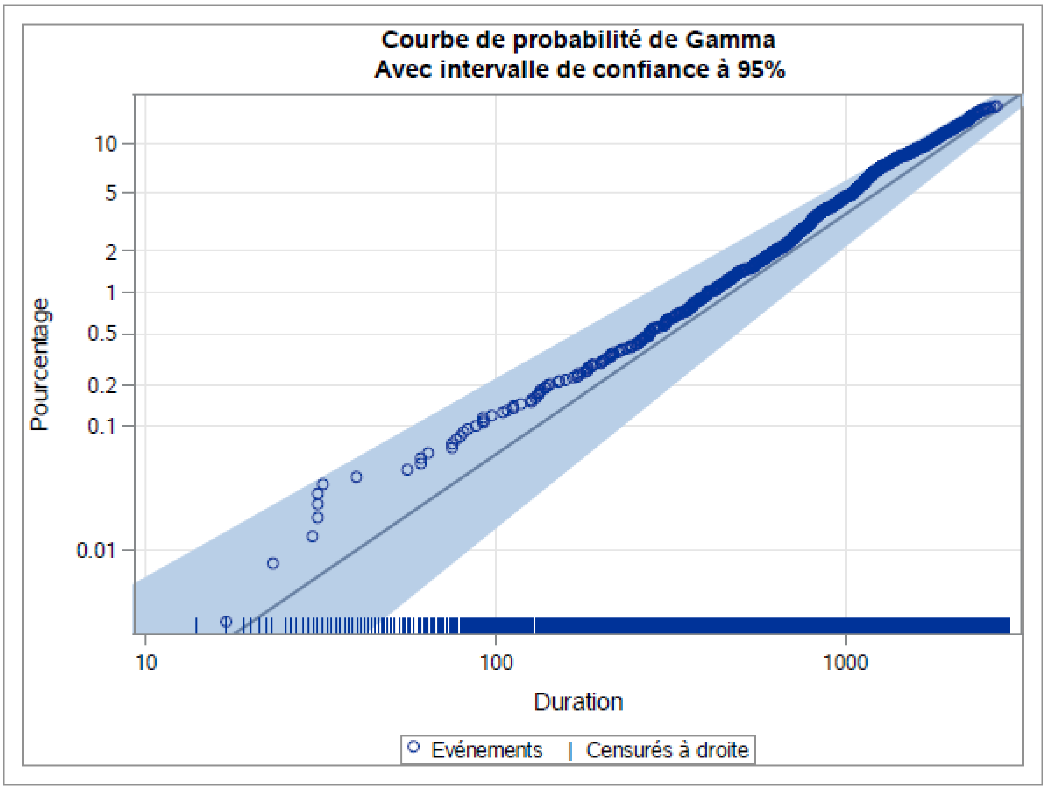


4. Appendix 4: Probability plot for each type of exit

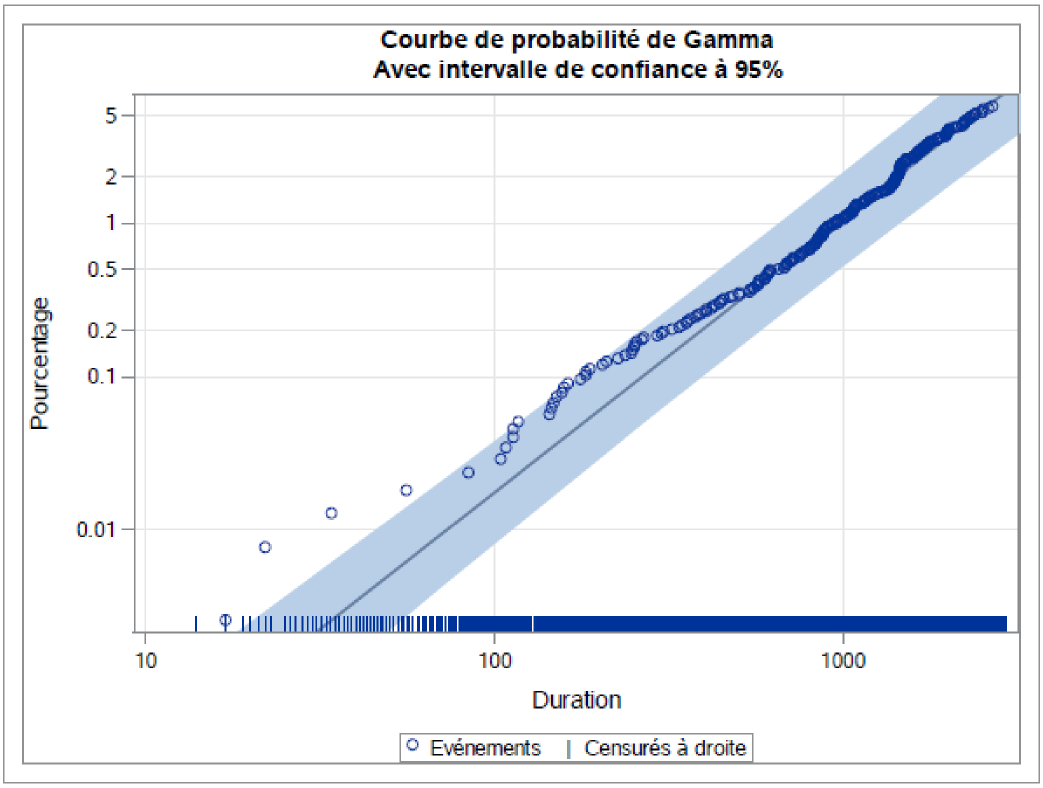
- Probability plot for acquisition exit



- Probability plot for liquidation exit



- Probability plot for IPO exit



5. Appendix 5: estimated results for the Cox proportional hazard model.

Appendix 5: Estimation results

Cox Proportional Hazard model

Coefficient	Acquisition		Liquidation		IPO	
	Estimate	Hazard Ratio	Estimate	Hazard Ratio	Estimate	Hazard Ratio
Early	0,24947 ***	**28%**	-0,4667 ***	**-37%**	1,04413 ***	**184%**
Expansion	0,23557 ***	**27%**	-0,39595 ***	**-33%**	1,22617 ***	**241%**
Later	0,18775 ***	**21%**	-0,42724 ***	**-35%**	2,03391 **	**664%**
roundNB	0,19796 ***	**22%**	0,00434	**0%**	0,45634 ***	**58%**
Amount	0 **	**0%**	0 ***	**0%**	0 ***	**0%**
Participants	0,04186 *	**4%**	-0,07088 **	**-7%**	0,0726 ***	**8%**
Angel_dummy	0,1959 **	**22%**	0,05673	**6%**	-0,56092 ***	**-43%**
VC_dummy	0,33353 ***	**40%**	-0,03386	**-3%**	-0,89824 ***	**-59%**
Milestones_round	0,03525 ***	**4%**	-0,05863	**-6%**	0,07385 **	**8%**
advertising	-0,11948 ***	**-11%**	-0,82849 ***	**-56%**	0,86486	**137%**
ecommerce	-0,57921	**-44%**	-0,31262 **	**-27%**	0,21597 ***	**24%**
social	-0,14219 ***	**-13%**	-0,58173 ***	**-44%**	0,97288	**165%**
software	-0,26602	**-23%**	-0,70224 ***	**-50%**	0,32443 ***	**38%**
hardware	-0,7692 *	**-54%**	-0,29505 *	**-26%**	0,55135 ***	**74%**
mobile	0,02074	**2%**	-0,36117 ***	**-30%**	-0,18772	**-17%**
enterprise	-0,07959	**-8%**	-1,073320 ***	**-66%**	0,10871	**11%**
network_hosting	0,1984 ***	**22%**	-0,3497 *	**-30%**	0,84434 *	**133%**
service	-0,34831 ***	**-29%**	-0,8926 ***	**-59%**	0,73505 ***	**109%**
education	-0,45652 ***	**-37%**	-1,934850 ***	**-86%**	1,61455 **	**403%**
games_video	0,01888	**2%**	-0,03857	**-4%**	-0,22392	**-20%**
IPO_market_global	0,0002007	**0%**	-0,00604 ***	**-1%**	-0,00468	**0%**
IPO_market_tech	-0,00124 *	**0%**	-0,00997 **	**1%**	-0,01226	**-1%**

Estimated coefficients for the Cox proportional hazard model set in the framework of competing risks with three types of exit: acquisition, liquidation, and IPO. The duration is defined as the number of days between the start of a given financing round and the time of the exit. In the case of non-exited rounds, their duration is considered as right-censored at the date of the 31st December 2013 and included in the model. A ***, ** and * indicates that the coefficient is significant at a 1%, 5% and 10% level respectively.