Corporate venture capital: The impact of industry relatedness between CVC parent company and startup on the startup's performance

Marcel Aksoy

152118087

Dissertation written under the supervision of Alberta Di Giuli, PhD (ESCP Business School)

Dissertation submitted in partial fulfilment of requirements for the MSc in the International Master in Management with Major in Strategy and Consulting (Double Degree with ESCP Business School), at the Universidade Católica Portuguesa, 02.06.2020

Corporate venture capital: The impact of industry relatedness between CVC parent company and startup on the startup's performance

Marcel Aksoy
Universidade Católica Portuguesa, MSc in the International Master in Management with Major in Strategy and Consulting (Double Degree with ESCP Business School)

Abstract

M&A literature reports that the relatedness of an acquiror and a target positively impacts the outcome of the deal and the subsequent performance of the new entity. The superior performance is explained by resource-based and transaction-cost arguments. In corporate venture capital (CVC), investors and startups are reported to benefit both from a "symbiotic relationship" when the CVC's investment is driven by strategic intentions. These relationships result in more positive outcomes for both the startup and the CVC, measured among others by startup valuations, innovation output or financial performance. Based upon these findings, this study explores on whether industry relatedness between a CVC and a startup impacts the startup performance. Using a dataset of 891 CVC deals from European CVC investors it investigates how industry relatedness impacts the occurrence of a startup IPO. The results of the analysis suggest that no significant relationship exists between industry relatedness and the IPO likelihood of the startup. However, this study is subject to several limitations that leave space for future research in this domain.

Resumo

A literatura sobre fusões e aquisições relata que o relacionamento de um comprador e de uma empresa-alvo tem um impacto positivo no resultado do negócio e no desempenho subsequente da nova entidade. O desempenho superior é explicado por argumentos baseados nos recursos e nos custos de transacção. No Corporate Venture Capital (CVC), os investidores e as startups beneficiam ambos de uma "relação simbiótica" quando o investimento do CVC é motivado por intenções estratégicas. Estas relações resultam em resultados mais positivos tanto para a startup como para o CVC, medidos, entre outros, por valorização de startups, produção de inovação ou desempenho financeiro. Com base nestes resultados, este estudo explora se a relação em termos de indústria entre um CVC e uma startup tem impacto no desempenho da startup. Utilizando uma amostra de 891 negócios de CVC de investidores europeus, este projeto investiga de que forma a relação entre a indústria de cada uma das partes tem impacto na ocorrência de uma OPI pela startup. Os resultados da análise sugerem que não existe uma relação significativa entre a relação entre a indústria e a probabilidade de uma OPI de startup. No entanto, este estudo está sujeito a várias limitações que deixam espaço para a investigação futura neste domínio.

KEYWORDS

corporate venture capital, industry relatedness, startup performance, deal performance

Table of content

List of tables

1. Introduction

Corporate venture capital (CVC) is evolving into an increasingly used vehicle for large corporations worldwide to invest in innovative startups and to harness financial returns as well as to access innovative, potentially disrupting technologies. Even though the model of CVC is not new in corporate literature and companies such as Microsoft, Intel or Xerox invested largely through CVC around the turn of the millennium, the industry recently sees immense growth with CVC investors coming also from non-tech sectors such as fashion, retail or consumer goods. This is underlined by the number of CVC deals, which reached 429 in the fourth quarter of 2018, constituting a 59% year-over-year growth. The notion that CVC is an increasingly important item on CEOs' agendas is further supported by the fact that 264 new CVCs invested for the first time in 2018, representing a 35% year-over-year growth (CB Insights, 2019). Considering the recent and ongoing relevance of CVC in the corporate investment context, this study aims at providing further evidence on the success-defining mechanisms of CVC deals.

A large part of CVC literature has so far concentrated on comparing CVC investments with those of regular VCs. The most compelling difference is thereby the strategic nature of CVC investments. While regular VCs in general exclusively focus on return on investments, CVCs are further interested in gaining access to a startup's resources, technologies or products. On the other hand, the CVC supports the startup, e.g. with its process know-how, customer base or own existent products (Dushnitsky & Lenox, 2006; Colombo & Murtino, 2017). Studies report that CVC investments with strategic intentions have a higher likelihood of a successful exit compared to financially driven investments (Gompers & Lerner, 2000a), have better post-IPO (Initial Public Offering) long-run stock returns (Chemmanur et al., 2012) and obtain higher startup IPO-valuations than startups backed by VCs (Ivanov & Xie, 2010).

A common reason why CVCs engage in strategic startup investments are the technologies or patents of a startup. Hence, CVC is often termed as a window to new technologies (Maula et al., 2013) and a means for corporate investors to foresee potential disruptors and, ideally, to diminish the risk of being disrupted by acquiring shares in those new companies, thus gaining access to their technology or products. But when doing so, do CVCs really go beyond their industry boundaries to detect new technologies or potential competitors? Or do they stay within their sector, identifying potential disruptors that are close to their core business?

CVC literature indicates that the industry relatedness of the CVC and the target startup impacts the decision-making process prior the acquisition – both on the CVC as well as on the startup side. Dushnitsky and Shaver (2009) speak about the "paradox of corporate venture capital" when adhering to their research findings that startups with products that substitute those of a CVC's parent company are less likely to accept an investment by the latter due to fears of imitation and expropriation. The opposite effect of a higher CVC investment likelihood was reported when the products of the CVC and the startup have a complementary relationship. The literature reports also the impact of industry relatedness once an investment took place, mainly in regards to innovation. For example, startups backed by related CVC investments outperform unrelated ones in terms of patent output (Chemmanur et al., 2014).

So, when looking at CVC literature, we can assume that CVC investors have an advantage compared to regular VCs due to their strategic valued added to the startup. Also, the relatedness of the CVC and the startup seems to impact the likelihood of forming CVC investments in the first place and to influence the innovative capabilities of a startup. What remains, however, is the question on whether the strategic advantage of a related CVC investment also translates into an overall stronger business performance of the startup? The literature on industry relatedness provides insights by advocating that related acquisitions are more beneficial to both acquiror and target. Among others, it bases its argumentation on the related view which emphasizes the capacity of knowledge sharing between investor and target (Oxlex & Sampson, 2004), transaction-cost arguments (Villalonga & McGahan, 2005) or the resource relatedness between both parties (Chang et al., 2017). Nevertheless, the role of industry relatedness within CVC investments is still relatively unexplored. While studies such as Achleitner et al. (2014) use the concept of industry relatedness to explain trade sales returns of regular VCs when selling their shares in a startup to a strategic buyer, there is no further literature on how industry relatedness in a CVC-startup dyad impacts the performance of the startup.

Therefore, the aim of this work is to extend the existing CVC literature by adding the concept of industry relatedness as a driver of CVC investment performance and to analyze its role in impacting the success of a CVC investment. While most CVC research is conducted using US datasets, this study is using a sample of 891 CVC deals of CVC investors headquartered in Europe. Thereby it also adds to the further exploration of European CVC deal activities. Similar to prior research in CVC, this study uses the occurrence of a startup IPO as a measurement for startup performance and, consequently, for a CVC deal's success. The guiding research question hereby is:

Do startups that are funded by corporate venture capitalists (CVC) perform better when the CVC parent and the startup are in the same industry?

The outcome of this study indicates that the formation of a CVC-startup dyad relationship with an industry relatedness does not increase the likelihood of the startup to achieve an IPO compared to CVC-startup investments that are of unrelated nature. This result was supported by several robustness tests conducted such as using different industry classifiers or a subsample with exclusively listed CVCs. It thus does not confirm previous studies which report an impact of industry relatedness, measured by a SIC code proximity, such as Chang et al. (2017) who found that industry relatedness strongly predicts a firm's new product market entry. Also, M&A literature suggests different empirical results by reporting that industry relatedness influences acquirors' target preferences (Stellner, 2015) and, based on combined-resources and transaction costs theories, that acquirors most likely prefer acquisitions over other investment types if they share an industry relatedness with the target (Villalonga & McGahan, 2005). In VC literature, industry relatedness between a strategic buyer and the startup is reported to negatively influence the return of a VC who wants to sell its shares in the focal startup via a trade sale (Achleitner et al., 2014). In the context of CVC literature, the insignificant results of this study do not confirm findings like those of Ivanov and Xie (2010), who use a qualitative approach to measure industry relatedness, and who report higher startup IPO valuations and higher takeover premiums for startups when they exhibit an industry relatedness with the invested CVC. Other CVC researchers like Chemmanur et al. (2014) confirm the positive impact of industry relatedness in CVC dyads by stating that a startup's innovation performance is higher when the CVC investor is in the same industry.

The dataset of this study revealed a skewed distribution of SIC industries with a large concentration of deals that involved CVCs from chemical companies and pharmaceuticals. The skewness was even stronger when looking at the CVC investor industries which accounted for the sample's IPOs, showing that CVCs from chemical companies and pharmaceuticals were involved in 72,3% of the total startup IPOs. The notion that industry-affiliation, rather than industry relatedness, was the driver for a startup IPO was confirmed when adding industry control variables to the regression model which were highly statistically significant and at the same time rendered the coefficient of industry relatedness insignificant. Furthermore, this observation is in line with CVC literature reporting that CVCs that invest frequently in CVC investments exhibit higher cumulative abnormal returns compared to sporadic investors as they build up a reputation as a stable and trustworthy investment partner (Benson & Ziedonis, 2009). The strong representation of pharmaceuticals, chemical companies but also telecoms and media companies in the sample underlines that companies are more likely to invest in CVC activities when they operate in industries with rapid technological change and high competitive intensity (Basu et al., 2011). Park and Steensma (2012) report significant differences in the likelihood of a startup IPO between complementary CVC-startup relationships and substitutional ones that could not be confirmed within this study. However, their results indicate that this finding is contingent on the industry in which CVC and startup operate, which again is in line with this paper's observation of a significant effect of industry affiliation on the CVC investment performance.

Aside from industry control variables, the total funding of a startup was the only control variable that exhibited a highly statistically significant and economically relevant impact on the likelihood of a startup IPO. While the notion that startups that receive more funding tend to be more successful seems straightforward, it poses the issue of endogeneity when investigating success-determining factors of CVC investments. In particular, startups that already received high sums of funding might be more attractive for successful investors (e.g. pharmaceuticals), thus blurring measures that aim at studying the impact that investors can have on the startup's success. Among others, Masulis and Nahata (2009) and Park and Steensma (2012) reported endogeneity in CVC research and proposed measures to alleviate this issue that will be discussed within this work.

The paper commences with a literature review which classifies CVC in the context of corporate venturing activities, illustrates the concept of industry relatedness and its impact on CVC-startup relations, and describes the drivers for CVC investment performances, both independently and in the context of industry relatedness. Based upon the literature review, the hypothesis of the paper is presented, linking industry relatedness to the performance of the CVC investment. Following that, the dataset and the methodology are presented before showing the results of the analysis, including descriptive statistics, the main probit regression model and robustness tests. Finally, the results of the latter and the study's limitations will be discussed and, based on that, future research potentials suggested.

2. Theoretical background

Chapter 2 is a literature review that, first, provides an overview of corporate venture capital (CVC), secondly, presents the concept of industry relatedness and its role within CVC and, thirdly, further illustrates the performance drivers of CVC investments and how industry relatedness impacts these drivers.

2.1 Venture capital, corporate venturing and corporate venture capital

Venture capital (VC) firms are financial intermediaries that invest investors' money directly in portfolio companies (Metrick & Yasuda, 2011). As the investment targets are non-listed private companies, venture capital is characterized as a type of private equity investment which focuses on capital investments in young and innovative companies. Venture capital investments are normally minority share investments and come along with representation in the board of the invested company and management support (Schefczyk, 2001). Contrary to strategic investments of corporations, venture capital firms are driven by financial returns and seek an exit of their investment (Metrick & Yasuda, 2011).

In addition to the financial return driven purpose of venture capital firms, corporate venturing or corporate entrepreneurship activities mainly follow strategic goals such as the access to new innovations (Chesbrough, 2002). The literature uses the term *corporate venturing* as an umbrella term for all entrepreneurial activities, both internal and external. While internal corporate venturing represents investments and foundations of ventures within a company, external corporate venturing comprises activities like corporate venture capital (Dauderstaedt, 2013). Even though the existing literature agrees on the differentiation between *internal corporate venturing* as a means to promote opportunities within the firm and *external corporate venturing* as a focus on opportunities outside the firm in the form of startup investments, many corporate venturing units pursue some combination of internal and external venturing (Birkinshaw & Hill, 2005).

Corporate venture capital (CVC), according to Chesbrough (2002), describes the direct investment of corporate funds in external startup companies. In alignment with the definition provided by Dauderstaedt (2013), CVC does not include corporate venturing activities where the startup, even though operating separately from the core business, remains part of the legal entity of the parent. The definition of Chesbrough (2002) further excludes the investment of a company via an externally managed third-party fund even if the investment vehicle is exclusively funded by one investing company. Others, such as MacMillan et al. (2008) or Maula (2001), define CVC investments as direct capital investments in startups that can be done by third-party funds, dedicated funds or self-managed funds either on their own, alongside traditional VCs or in a syndicate of investors. This paper will further use the term *CVC* following the definition of Chesbrough (2002), however, not limiting it solely to self-managed corporate-funds but also including other investment vehicles as long as the corporate's major stake in the fund is apparent.

2.2 Industry relatedness in corporate venture capital investments

2.2.1 Definition and role of industry relatedness in investments

Industry relatedness is a commonly used concept within the acquisition framework. In that context it defines the degree to which the acquiring and the target firm are active in related markets which indicates that they share common characteristics such as the usage of similar or complementary resources, knowledge basis, technologies and products (Lim & Lee, 2016). Industry relatedness in acquiror's target selection and the resulting corporate coherence play a key role for the acquiror's strategy and its growth (Cefis & Rigamonti, 2018).

Previous works analyzing the role of industry relatedness in the context of mergers & acquisition, private equity or (corporate)venture capital ((C)VC) use different terms to describe the above-mentioned concept or similar concepts such as *market relatedness* (Chang et al., 2017), *product-market relationships* (Masulis et al., 2009), *substitute or complementary products* between CVC and startup (Dushnitsky & Shaver, 2005), *relational view on CVC* (Weber et al., 2016) or *strategic fit* between CVC parent and venture (Ivanov & Xie, 2010). An in-depth explanation and analysis of the different concepts is out of scope of this work and henceforth the definition of Lim and Lee (2016) will be used when speaking of *industry relatedness*.

The understanding of industry relatedness within the context of M&A is based upon the differentiation between related and unrelated diversification (Capron, 1999) with empirical evidence indicating that related diversification produces more positive outcomes. As the levels of expected returns from related acquisitions are higher than for unrelated ones, the chance of deal completion is accordingly higher for related acquisitions. In addition, related acquisitions exhibit a lower perceived risk due to a high level of knowledge and low information asymmetry between acquiror and target. Moreover, acquirors tend to pay a higher premium for related businesses (Gondhalekar et al., 2004) which in turn results in a higher acceptance of a deal by the target (Wong & O'Sullivan, 2001). In the context of cross-border acquisitions, related and strategically motivated take-over deals are more likely to be completed than unrelated and financially motivated deals (Lime & Lee, 2016). Villalonga and McGahan (2005) report that industry relatedness is associated with the choice of firms on how to enter in external acquisitions or partnerships. When acquiror firm and target firm exhibit an industry proximity, the target firm is more likely to prefer an acquisition over an alliance, and an alliance over a divestiture. Also, relatedness between the acquiror firm's industry and the activity that is subject to the transaction result in the preference of acquisitions over alliances, and alliance over divestitures. Both effects can be explained by a combination of resource-based and transaction-cost arguments that suggest that greater relatedness leads to lower cost of integration due to economies of scale (Villalonga & McGahan, 2005). Another argument for the preference of acquisitions is the related view which implies that industry relatedness and the direct competition between acquiror and target lead to more integrative and protective governance structures to enable knowledge sharing between the two firms (Oxley & Sampson, 2004).

Chang et al. (2017) provide an additional view on industry relatedness by linking a firm's entry into new product markets to its resource relatedness. Prior research decomposes a firm's resources into market-related and technology-related resources (Nerkar & Roberts, 2004; Sosa, 2009). While technological resources comprise upstream capabilities such as technological knowledge and skills, market-based resources are linked to the brand, existing customer relationships and the understanding of customer preferences. Based on the resource-based view, Chang et al. (2007) argue that market relatedness does predict market entry, while technological relatedness does not. This effect can be explained with the "stickiness" of market resources that often are non-transferrable and specialized within a firm, while technological resources are tradeable e.g. through licensing agreements. Another argument is that managerial attention is generally focused on downstream market-related resources in order to assess product-market opportunities or resource acquisition.

In the context of venture capital investments, Achleitner et al. (2014) report lower returns for regular VC firms when they sell their shares via a trade sale to a strategic acquiror that has synergetic capabilities. Accordingly, trade sales to strategic acquirors that do not yet operate in the same industry as the invested startup yield higher returns for the exiting VC firms.

2.2.2 Impact of industry relatedness on CVC investment decisions

The aim for strategic returns in addition to financial returns differentiates CVCs from traditional VCs. This strategic view can impact not only the investment decision of the CVC firm to acquire a startup but also the decision of the startup on whether to engage in a partnership with a CVC firm. One major strategic objective of CVCs is to increase the sales and profits of the corporate parent through its investments. Hence, CVC investment seeks to identify startups that offer the potential to exploit synergies between the two organizations. If strategic considerations overwhelm in the CVC decision-making process, CVCs accept a trade-off of lower direct returns from the startup investment as long as it leads to improved performance of the parent organization (Chesbrough, 2002).

However, the decision to enter in an inter-organizational partnership lies not exclusively with the CVC. The startups, mostly young firms with an innovative idea at the core of their business model, have also interests and fears that potentially lead to a rejection of a CVC investment. In this context, industry relatedness can have a major impact on both the CVC's and the startup's decision to enter in a dyad relationship. A startup's product can be a potential substitute and could render corporate products and services obsolete. In this case, the corporate investor has an incentive to behave opportunistically and to copy the startup's new technology (Gans et al., 2001) causing the parties to remain with two diametrically opposed points of view: A CVC firm that is unwilling to invest as long as the startup does not disclose its invention and a startup that rejects to do so because of fears of imitation and expropriation (Dushnitsky & Shaver, 2009). This "paradox of corporate venture capital" leads to the effect that CVC investments in startups with substituting products are less likely to materialize. The lower likelihood of startups to accept CVC funding in case of a substitutional relationship contradicts to the view of many corporations that perceive CVC activities as an early alert system and vehicle to acquire potentially disrupting inventions at an early stage. Also, it contradicts to empirical research indicating that strategic CVC-backed investments have a higher likelihood of a successful exit compared to financially driven investments (Gompers & Lerner, 2000a), have better post-IPO long-run stock returns (Chemmanur et al., 2012) and obtain higher startup IPO-valuations than startups backed by independent VCs (Ivanov & Xie, 2010). However, the expropriation fear and greater moral hazard concerns are also reflected in higher valuations extracted by a startup when it receives funding by a competitive CVC (Masulis & Nahata, 2009). On the other side, if the corporate products and the startup's invention have a complementary relationship, Dushnitsky and Shaver (2009) report a higher likelihood of forming a CVC-startup relationship. Here, the corporate parent has less incentive to misuse the disclosed information and can benefit from the complementarity as it may secures demand to its own products (Anand & Galetovic, 2004).

Once a CVC firm is already invested in a portfolio startup, the likelihood of an ultimate acquisition of the latter increases when CVC parent and startup operate in a similar industry (Dimitrova, 2015). Koehn (2018) argues that depending on a corporate parent's degree of explorative or exploitative capabilities, previous CVC investments in startups influence the likelihood of a subsequent entire acquisition of the startup. More specifically, adding the concept of industry relatedness, they state that CVCs with a high degree of exploitative orientation tend to acquire startups with a high product-market overlap in order to keep their competitive advantage and to strengthen their knowledge base (Siren et al., 2012). On the contrary, product-market overlap is negatively associated with an ultimate startup acquisition when the CVC has a high degree of explorative orientation.

2.2.3 Industry relatedness and CVC-startup relationship

Not only does industry relatedness influence the decision on whether or not to engage in a partnership, it also determines the collaboration-, strategy- and governance-decisions once a CVC-startup dyad is formed. The relationship between CVC and startup is sensitive to the startup's impact on the CVC parent's business as the CVC, under some conditions, might choose to act adversely to the startup's interest and success. Contrary to independent VCs, CVCs follow also strategic objectives and are mostly integrated within corporate structures next to business lines that are potentially sensitive to the startup's activities. This poses challenges to the establishment of stable relationships between CVC and startup unlike it is the case for independent VCs who, limited to their financial objectives, rely mostly on their reputation and hence achieve greater alignment between them and the startup (Sahlmann, 1990). While an industry relatedness between CVC and startup can be of benefit for both parties as the CVC can add important non-financial backing, it can also turn out detrimental for the startup as the CVC parent is not only interested in the startup's profits but its total corporate profits (Dushnitsky & Shaver, 2009).

An advantage of CVCs that share the same industry with the startup is the provision of access to its current operational capabilities. This allows the startup to use the corporate's manufacturing plants, distribution channels, technology or brands and to adopt the corporate's practices to build, sell or service its products. However, in fast changing business environments with potential disruption, relying on existing capabilities might be a liability rather than a competitive advantage. In such a case, a CVC investment in an external startup offers the corporate the opportunity to develop different capabilities externally and, if successful, to integrate them into the core organization (Chesbrough, 2002). The valued-added through CVC investors that follow strategic intentions is backed by research reporting that CVC-backed startups are able to obtain higher valuations at the IPO and higher takeover premiums than startups backed by independent VCs. However, this superior performance in CVC investment compared to independent VCs is only observed when the CVC has a strategic focus (Ivanov & Xie, 2010).

Industry relatedness further impacts financial contracting decisions in CVC-startup relationships that address the potential conflicts of interest of a CVC investment. Masulis and

Nahata (2009) distinguished between competitive and complementary CVC investors based upon the degree of overlap in SIC codes between CVC parent and startup. They report that startup insiders, such as founders, award lower board representation to competitive CVC investors compared to complementary investors. Also, startup insiders retain greater board power when competitive CVCs are invested. The startup's desire to limit the influence of CVCs is further reflected in the result that CVCs who are lead investors attain less board representation than independent VCs who act as lead investors (Masulis & Nahata, 2009).

2.3 Performance of CVC investments

2.3.1 Drivers of performance in CVC investments

Prior research has shown that financial returns of CVC funds tend to be more volatile compared to those of the general venture capital market. Dushnitsky and Lenox (2006) explain this variation in CVC performance with the different orientations and objectives between CVC funds. In particular, strategic considerations and the derived, indirect, benefits from the investment may overwhelm for some CVC funds. Strategic benefits can include a window to new technologies and the connection to innovative startups provides the possibility to access those technologies through licensing or acquisitions. In this context the complementarity of products and services between CVC parent and startup can enable the increase in demand for current or future corporate products. Taking these factors into account, firms that pursue CVC activities primarily as a window to new technology create more firm value than CVCs that mainly seek financial returns (Dushnitsky & Lenox, 2006). One of the main objectives of a new startup and its investors is rapid growth, often reflected in growth in sales. While this objective is common among CVCs and independent VCs, the channels to reach it differ between these two types of investors. Independent VC investments come along with a higher increase in headcount, payroll expenses and fixed assets compared to CVC investments. Colombo and Murtino (2017) attribute this difference to the symbiotic relationship between corporate and startup in which the startup benefits from corporate resources and does not need to hire substantially more employees. Other findings show that independent VCs are superior in helping startups transform their ideas into viable companies through their skills in developing strategy, hiring key employees or obtaining additional financing. CVCs, on the other hand, provide startups with commercial and public credibility through their widely known brands, thus helping them attracting customers, suppliers and partners. When entering a partnership with a technology- and research-focused CVC investor, startups can further benefit from those resources by receiving support for their product development (Maula & Murray, 2001).

Firms that invest more in internal innovation activities also gain higher returns from CVC activities. In particular, the realization of an acquisition benefit for the CVC parent is related to the degree of internal innovation activities compared to CVC investment activities. Previous findings show that as investments in CVC activities increase in magnitude compared to the CVC parent's R&D expenditures, the beneficial effects of CVC activities on acquisition performance diminish (Benson & Ziedonis, 2009). Companies with a strong internal knowledge basis exhibit a higher likelihood to leverage technologies obtained through investments, a capacity that is often referred to as "absorptive capacity" (Cohen & Levinthal, 1990). Another factor that is positively related with the acquisition performance of the CVC is the frequency of CVC investments. CVC firms with more stable CVC programs reported higher cumulative abnormal returns than investors that invested more sporadically. An interpretation of the positive effects of frequent CVC investments is that CVC firms with constant investments build up a reputation as a stable and trustworthy investment partner. Establishing such a reputation can have a strong impact on the access to potential target startups, especially when considering expropriation risks that cause many startups to be reluctant when it comes to corporate investments. Thus, frequent CVC investments can positively impact the relationship between the CVC and the startup as well as the performance of the CVC investment (Benson & Ziedonis, 2009).

Literature shows that the success of CVC investments can be contingent on the industry of either the CVC parent or the startup. Dushnitsky and Lenox (2006) report that the contribution of CVC activities to CVC parent firm value was highest in the devices and information technology sector. The notion that specific industries gain more value from CVC investment is supported by the differences in probability at which sectors engage in CVC. Specifically, firms that operate in industries with rapid technological change, high competitive intensity and weak appropriability engage more in CVC activities (Basu et al., 2011).

Another factor influencing the performance of CVC activities is the structure of the CVC investor. Previous literature argues that CVC units that enjoy full authority over their investment portfolio and that operate autonomously perform better as they can act more aggressively in their investment choices and are free from corporate resource restrictions (Birkinshaw & Hill, 2005; Yang et al., 2016). Drawing upon the exploration and exploitation framework (March, 1991), Lee at al. (2018) state that the structural autonomy of the CVC unit is positively related to the CVC parent's explorative innovation performance, while it is negatively related with its exploitative innovation performance.

Acting at the intersection between traditional venture capital and internal corporate investment, CVC firms can struggle in finding the balance between those two approaches which differ strongly, e.g. in terms of portfolio evaluation or management compensation (Fleßner et al., 2019). So instead of successfully creating additional value by leveraging both approaches, CVC firms end up weakening each of them, resulting in a “stuck in the middle” syndrome. This dilemma can be avoided by facilitating the knowledge transfer between corporate and startup by incentivizing corporate employees to share their knowledge as well as through the creation of communication protocols that enable the CVC unit to act as a middleman between corporate management, corporate employees and the startup, e.g. through its involvement in corporate strategy discussions (Fleßner et al., 2019).

2.3.2 Impact of industry relatedness on CVC investment performance

The existing literature exhibits several indications that a superior performance of CVCs compared to independent VCs can be attributed towards a fit in market, product or technology between the CVC firm and the startup. CVC-backed startups tend to be more innovative measured by their patenting outcome, although they are younger, riskier and less profitable than startups backed by independent VCs. The ability to nurture innovation is not only superior for CVCs compared to independent VCs, but also for CVCs that exhibit an industry relatedness with the startup compared to CVCs that invested in unrelated startups. The first explanation for this outperformance is the technological fit between CVC and startup, while the second one is the greater failure tolerance of CVCs compared to independent VCs (Chemmanur et al., 2014).

Drawing upon the differentiation between a competitive or complementary relation of the CVC and the startup, research indicates that a CVC deal’s performance is stronger for a complementary relationship rather than for a substitutional one. Assessing the performance of startups in terms of the ability of the CVC firm to exit through an IPO, Park and Steensma (2012) state that CVC funding is particularly beneficial for startups when they require specialized complementary assets or operate in uncertain environments. When specialized complementary assets are needed (e.g. hardware manufacturers may require customized software for their services), transactions costs in economic exchanges as well as the high costs of internal development pose challenges for new startups. Equity funding through an CVC investor with complementary assets can therefore mitigate these challenges. Accordingly, the positive impact of CVC investments is stronger for startups that seek complementary assets than for startups that need generic assets (Park & Steensma, 2012).

2.4 Hypothesis

As illustrated in the previous paragraphs, the relatedness of an investor and the associated startup can produce more positive outcomes for both investor and startup than in non-related investment dyads. Based upon resource-based and transaction-cost arguments, existing research indicates that a strategic investor can share valuable knowledge, resources or supplier- and customer-relationships with a startup aside from providing financial funding. The literature on industry relatedness and strategic CVC investments indicates that this effect is even stronger when both parties share the same industry or products. This access to corporate resources in turn should increase the likelihood of a startup to succeed while at the same time the corporate investor benefits through the access to new technologies, customers or market segments. Hence, I hypothesize that a CVC-startup investment relationship will be more beneficial to the startup's performance if both are from the same industry:

> *Hypothesis: CVC funding will be more beneficial to the startup's performance if the CVC parent company and the startup are in the same industry.*

3. Data and methodology

This chapter starts with the sample and its data sources. In the second part it presents the different measures for CVC investment performance used in literature before describing the dependent, independent and control variables applied in the analysis. It concludes with a specification of the econometric model.

3.1 Sample and data sources

This study uses deal data from the Private Equity section of Thomson OneBanker (formerly VentureXperts) as main source of information on new startups, startup IPO dates, CVC investors and CVC funding deals including SIC codes of startups and CVCs to define the industry relatedness in the dyad relationships. Thomson OneBanker further provides comprehensive information for (C)VC deals including deal sums, age of startups, information on investment rounds and composition of the investors syndicate. Considering the analysis of European CVC deals, the extensive availability of European deal data was deemed an important factor in selecting Thomson OneBanker.

The initially retrieved dataset from Thomson OneBanker included all deals from CVC investors with headquarters in Europe from 01.01.2002 until 31.12.2014. The latter date was selected in order to be able to measure the effect of industry relatedness on a startup's IPO as most (C)VC investors exhibit holding periods of several years before they exit via a startup IPO. It included all European corporate private equity and CVC deals in the given time period while excluding real estate investments, resulting in 2102 initial dyad observations between a CVC investor and a startup. It is not unusual that CVC investors invest several times in the same startup over the course of different funding rounds. Hence, to properly measure the outcome of each unique dyad relationship and in order to avoid redundant measurements, only the last investment of a CVC investor in a startup was taken into the dataset. Valuable information from the removed observations of earlier rounds, such as the number of the first participated funding round of a CVC, were maintained in the dataset by adding a new column containing the information. This step decreased the observations to 1373 CVC deals.

In a next step, all CVC investors in the dataset that were not owned or did not exhibit a major affiliation to a corporate parent were removed as well as those whose parents were financial investment companies or state-owned agencies. Furthermore, Thomson OneBanker includes deals that are missing data, e.g. for the invested deal sum of a CVC, a problem which is also reported by previous works such as Colombo and Murtino (2017) and Ivanov et al. (2010). Accordingly, those deals with missing observations were also removed from the dataset, resulting in a final dataset comprising 891 deal observations.

To construct the main independent variable, SIC codes were retrieved from Thomson OneBanker. In some cases, SIC codes for either one or both CVC investor and startup were missing. Here, the missing SIC codes were added manually by conducting online research about the nature of the business and by using the publicly available database "siccode.com". Eventually, of the sample size of 891 observations 210 deals exhibit a CVC-startup relationship that is classified as related based on a 2-digit SIC code match.

For the dependent dichotomous variable *startup IPO*, data could be retrieved directly from Thomson OneBanker. Of the 891 deals 101 resulted in a startup IPO.

3.2 Methodology

3.2.1 Measures of performance of CVC investments

Similar to regular VC firms, CVC investors seek financial returns through exits such as IPOs or trade sales to third-parties (Gompers & Lerner, 2000a). Measuring financial returns, Cochrane (2005) uses the expected return, standard deviation, alpha and beta of regular VC investments. Another commonly used measure in VC literature (Achleitner et al., 2015;

Cochrane, 2005) is the Internal Rate of Return (IRR) from the perspective of the VC firm. The IRR is calculated on a deal basis and depicts a discount rate of cash outflows and inflows from the portfolio company that sets the net present value to zero. However, the possibility of using the IRR in C(VC) research is constraint by the access to data with deal sums not being necessarily disclosed, especially when (C)VC investors exit through a trade sale.

An increasingly stronger body in CVC literature accounts for the strategic dimensions of CVC investment performance (Benson & Ziedonis, 2006; Dushnitsky & Lenox, 2005) as well as the different perspectives on a deal's success by extending the scope to the startup's performance (Chemmanur et al., 2014).

Measuring the CVC investment performance on the CVC investor side, Benson and Ziedonis (2006) use an event study of the stock market's reaction to the CVC acquisition announcement to estimate the discounted future value, net of purchase price from the acquisition. Dushnitsky and Lenox (2006) adopt Tobin's *q* as a measure of firm value to assess the success of CVC investments. According to the authors, Tobin's *q* is a good proxy for a firm's competitive advantage and, unlike measures such as the return on investment (ROI), accounts for risks and is less liable to reporting distortions.

Further expanding the performance measurement to strategic dimensions, a large body of literature focuses on the innovativeness and technological performance of the CVC investor's parent. While earlier research uses R&D expenditures as a proxy for firms' innovation activities, recent studies use patent-based metrics (Wadhwa et al., 2015; Belderbos et al., 2018) such as the number of patents and the number of citations received by patent as these metrics, how Chemmanur et al. (2013) argue, capture the actual innovation output and how effectively a firm has used its innovation inputs.

The measurement of performance of CVC investments in the existing literature is not restricted to the financial or strategic returns on the CVC acquirer side but further explores the impact on startups when they engage in a CVC dyad relationship. For example, Maula and Murray (2001) sent a questionnaire to 200 CEOs and founders of US high-tech startups asking them about the added- value to their firm by both independent VCs and CVCs. Colombo and Murtino (2017) conducted an analysis using the General Methods of Moments methodology to measure the acquired firm's overall economic performance with the logarithm of real sales value as output variable and the logarithms of real payroll expenses and real fixed assets as input variables. Evaluating the impact of CVC investment on a startup's innovation performance, Chemmanur et al. (2014) and Lahr and Mina (2016) used the startup's patenting output as performance measure.

3.2.2 Dependent variable: Startup IPO

This study uses the binary variable of a successful *IPO* of the invested startup as the dependent variable to measure the CVC investment's performance. Among the exit strategies of VC and CVC investors, an IPO or a trade sale belong to the most utilized forms (Achleitner et al., 2014; Gompers & Lerner, 2000a). As venture capital is by definition the investment in private, non-listed firms, deal sums and other investment information are not necessarily subject to public disclosure. Hence, the large part of VC and CVC research uses publicly available data such as from startup IPOs or CVC investors' annual reports while only a few use proprietary data to analyze trade sale performances (Achleitner et al., 2014). Whether or not an investment resulted in an IPO is a commonly used proxy to measure the success of a CVC investment and can be of advantage for both CVC investors and startups alike (Gompers & Lerner, 2000a). Comparing VC to CVC investments, Guo et al. (2015) for example report that CVC-backed firms receive larger investments than VC-backed firms and that a larger investment increases the likelihood of an IPO. Using a dataset of firms from the high-tech sector, Park and Steensma (2012) found that CVC funding increases the likelihood of an IPO when the startups require specialized complementary assets or operate in uncertain environments. From an investor's point of view, empirical data shows that exits via IPOs result in higher return on investments compared to exits via a trade sale (Park & Steensma, 2012). From the entrepreneurial perspective, the relative attractiveness of an IPO might be even more salient. This is due to the convention that VC investors normally receive preferred shares granting them, in case of a trade sale, the right to receive at least a payoff equal to their initial investment before other investors or founders get paid out (Kaplan & Strömberg, 2003). In case of an IPO, however, those preferred shares need to be converted into common shares, thus explaining a higher appeal of an IPO to the entrepreneur (Park & Steensma, 2012). Park and Steensma (2012) further argue that many acquisitions during the dot-com bust period of 2001-2003 were actually failures and that therefore trade sales acquisitions might not be classified as a successful outcome. This view is supported by studies that observed that average trade sales acquisition prices in the pre-dot-com bust period were $190 million in the high-tech industry but only $49 million after 2000 (Waguespack & Fleming, 2009). Following the above-mentioned arguments this paper uses the startup IPO as a proxy to measure the successful outcome of a CVC deal given the positive impact on both startup and CVC investors reported in the literature. It thereby counts all startup IPOs that took place between January 2002 and March 2020.

3.2.3 Main independent variable: Industry relatedness

The main independent variable of this study is *industry relatedness*, measured through SIC-Code proximity on firm level between the invested startup and CVC parent company. As documented in paragraph 2.2.1, the concept of industry relatedness is a widely applied approach to measure effects in M&A or venture capital. The methodologies to measure industry relatedness in the existing literature, however, differ.

The Standard Industry Classification (SIC) code-based methodology applied in this study is a commonly utilized measure for industry relatedness in M&A and CVC research (Achleitner et al., 2015; Chang et al., 2017; Villalonga & McGahan, 2005). It consists of a 4-digit code with a hierarchical nested structure in which each level is represented by a digit (Cefis & Rigamonti, 2018). The first two digits represent the major industry sector to which a business belongs whereas the third and fourth digits indicate the sub-classification and the specialization, respectively (UK Office for National Statistics, 2007). When comparing the SIC codes of two firms, a SIC code matching at the 2-digit and 3-digit level generally indicates a strong relatedness (Achleitner et al., 2015). In the context of a trade sale, however, Achleitner et al. (2015) define a horizontal integration between a strategic buyer and a startup only when both companies share the identical 4-digit SIC code. Other scholars (Stellner, 2015; Valentini & Di Guardo, 2012; Villalonga & McGahan, 2005) did not use a binary approach to classify relatedness SIC-based but set the variable to zero if the 2-digit SIC codes did not overlap, to one if the acquiror and the startup shared the same 2-digit SIC code but not the 3-digit SIC codes, equal to two if both parties shared the 3-digit SIC code but not the 4-digit SIC codes, and to three if they shared the exact same 4-digit SIC code.

The regression analysis of this study will take the SIC code-based relatedness measure as a binary variable on deal level setting it equal to zero if there is not an overlap of the 2-digit SIC codes between CVC parent firm and startup and to one if both firms share the same 2-digit SIC codes following Achleitner et al. (2015) and Capron (1999) that a strong relatedness can be assumed on the 2-digit SIC code level.

One shortcoming of the SIC-based approach is that it neglects the conditions under which firms combine resources to create value and that it leaves significant strategic relationships among industries unclear (Cefis & Rigamonti, 2018). A measure of industry relatedness that aims to alleviate these issues is the co-occurrence analysis. This approach is based on the survivor principle, first proposed by Teece et al. (1994), which implies that due to economic competition only those market participants survive that exhibit the most efficient mix of activities (Lien & Klein, 2013). In the existing literature on diversification, the co-occurrence analysis has shown its ability to outperform SIC-based measures, for example for predicting firms' diversification decisions (Lien & Klein, 2009). Cefis and Rigamonti (2018) extend co-occurrence and survivor principle-based relatedness measures to fit the acquisitions framework. They propose a unidirectional index that can be applied as an independent variable in econometric models. Building upon the co-occurrence method but adding directional flows to the observed firm level relationship, the index can be particularly useful for dyad acquiror-target relationships or other settings where the direction of flows is essential.

Other scholars like Cassiman et al. (2003) pursued a qualitative approach to assess the industry relatedness of the two deal entities. Conducting an in-depth analysis of 31 M&A deals, they assessed the relatedness by directly asking the parties involved in the deal with a questionnaire. When startups decide to go public, they are obliged to disclose business-related information and the corresponding IPO prospectus can offer valuable insights on a firm's fields of activities and its business strategy. Prior research has utilized the disclosed information to gain insights on industry relatedness between startup and acquiror. Ivanov and Xie (2010) for example searched for mentions in the IPO prospectus of the startup on whether there are any strategic relationships or alliances between startup and CVC parent.

Masulis and Nahata (2009) note that SIC code-based measures only provide a relative broad description of the relatedness between two firms and that startups are generally concentrated in a small number of SIC codes. A large body in CVC literature therefore uses the CorpTech Directory (Dushnitsky & Shaver, 2009; Masulis & Nahata, 2009; Chang et al., 2017) to classify industry relatedness. The pre-dominant argument to use the CorpTech Directory is its finer and more detailed classification of industries and product markets for companies, including startups, compared to SIC-based measurements. For example, a company that is allocated to a single 4-digit SIC code might be allocated to several product categories based on the CorpTech Directory, thus accounting for the more diverse nature of firms' product portfolios (Masulis & Nahata, 2009). However, the CorpTech Directory predominantly covers US companies and hence its database does not fit to this study's sample of European CVC investors.

This study will use the SIC code-based methodology due to its wide acceptance and application in literature, practical purposes of accessibility as well as resource limitations. It, however, acknowledges the limitations of this methodology and emphasizes the potential for future research in this field using different ways of measuring industry relatedness.

3.2.4 Control variables

The study applies a number of control variables that can potentially influence the performance of a startup. The application of a commonly used control variable accounting for investors' size through measures such as market capitalization or total assets value is not feasible with the study's dataset due to a large number of non-listed CVC investors that do not disclose these financial indicators. Hence, the dichotomous variable *CVC listed* is included to account for this issue.

As more successful startups are more likely to receive higher funding by investors, the *total funding of the startup* is added as a control variable with the natural log of the measure to reduce the skewness of the model and to increase normality. Also, successful startups more likely attract a higher number of investors. Therefore, the natural log of *total number of funds* invested in the startup is also controlled for in the model. Following the same line of reasoning, startups with more promising prospects potentially exhibit higher investments per individual investor. Hence, the *total equity invested by the CVC in startup* is incorporated into the regression with the natural log of the variable.

In line with Ivanov et al. (2010) and Park and Steensma (2012) controls for the *round number of the last investment* of a CVC-startup dyad and the *round number of the earliest investment* of the CVC in the startup were added. Further, the *age of the startup at the last investment (in months)* is included in the model. Again, the natural log was taken for all three variables.

On the CVC level, following Masulis and Nahata (2009), the dichotomous variable *CVC lead investor* is included. It reflects the nurturing capabilities of corporate investors that help startups and potentially positively impact the startup's success. In Masulis and Nahata (2009) the variable takes the value of 1 if the CVC investor participated in the first investment round and invested the largest amount in the startup across all funding rounds. Due to limited access to information and in scope with this work, the *CVC lead investor* variable will take the value of 1 if the CVC investor participated in the first investment round and, different to Masulis and Nahata (2009), if it accounts for at least 50% of the total funding of the startup. At the CVC level, with the natural log of *CVC total deals number*, another variable is integrated to control for the experience of a CVC investor as more active and experienced CVC investors are more likely to add value to startups (Benson & Ziedonis, 2009).

Lastly, it controls for *year of investment* of the CVC in the startup (in case of several investments the year of the last investment is taken) and *industry* fixed effects (based on 2-digit SIC codes) by using year and industry dummy variables.

3.2.5 Specification of econometric model

After having explained the sample of this study and the proxies used as dependent and independent variables we can rephrase the hypothesis of this paper to specify the econometric model:

> *A CVC investment is more likely to result in a startup IPO when the CVC investor and the target startup share the same 2-digit SIC code.*

The dependent variable of this paper is the occurrence of a startup IPO, taking the value of 1 in case of an occurrence of an IPO and 0 in case of no IPO. Due to the dichotomous nature of the dependent variable, following the approach of Park and Steensma (2012), this study applies a probit regression model. According to Wooldridge (2010) the probit model can be derived from the binary response model, an index model that takes

$P(y = 1|\mathbf{x}) = G(\mathbf{x}\beta) = p(\mathbf{x})$, where $0 < G(z) < 1$

The probit model then constitutes a special case of the above equation

$G(z) = \Phi(z) = \int_{-\infty}^{z} \phi(v)\mathrm{d}v$,

where $\phi(z)$ is the standard normal density

$\phi(z) = (2\pi)^{-1/2} \exp(-z^2/2)$.

Hence, the probit function of this study would be

$\Pr(startup\_IPO_i = 1|related) = \alpha + \beta_1\, related + \beta_2\, X_2 + \cdots + \beta_i X_i + \mu_i$,

where the effect of industry relatedness is measured upon the probability of the startup realizing an IPO. Here, however, the study faces the issue of endogeneity, a common phenomenon in CVC literature when measuring a variable's effect on a CVC deal or on the subsequent performance of the CVC acquiror or the target startup. Among others, Gompers and Lerner (2001b) reported this issue when measuring the impact of VC capital inflows on VC startup valuations. Lahr and Mina (2016) observed endogeneity issues when drawing inferences on how a VC investment impacts the patent output of the startup and Colombo and Murtino (2017) when comparing the impact of CVC and independent VC investments on the startup's overall economic performance.

Concerning this study, endogeneity could impact the validity of the results of the analysis as the effect of industry relatedness between CVC parent and startup on the likelihood of success of the CVC investment (proxied by the IPO of the startup) might not be unidirectional. Moreover, the prospect of a successful startup investment could be more likely in specific industries (e.g. pharmaceuticals sector), thus attracting particularly pharmaceuticals as CVC investors and making startups more likely to accept funding by those investors from the same industry in the first place.

Park and Steensma (2012), who study the impact of the dichotomous variable "CVC investment" on the "likelihood of an IPO of the invested startup", report no endogeneity due to self-selection as the startups are less likely to choose to pursue an IPO during the initial phase of raising CVC funds. However, according to the authors, resource needs and environmental conditions can still influence startups to self-select into raising CVC funding. To alleviate the problem of endogeneity, they propose a bivariate probit model as it is deemed appropriate for obtaining causal inferences in case both the dependent and the main independent variable are of dichotomous nature (Park & Steensma, 2012; Wooldridge, 2010). To isolate the variation in the predictor variable X which is not explained by the outcome variable Y they introduce a first equation with the treatment effect (with Y as the dependent variable and X as the main independent variable) and a second equation that estimates the selection effect with X as the dependent variable. Besides other control variables, the latter contains an instrumental variable Z that is highly correlated with the endogenous variable X and not correlated with the outcome variable Y of the treatment effect equation. Applying the methodology to this paper, an instrumental variable Z would be highly correlated to the formation of a related CVC-startup deal relationship but uncorrelated to the IPO of the startup. Table 5 in chapter 4.2 shows the pairwise correlations of all variables for which data was gathered in this study. The variable that exhibits the strongest correlation with the main independent variable *related* is the natural log of *CVC total deal number* ($r = 0{,}1144$; $p < 0{,}001$). However, as this variable is even stronger related to the dependent variable *startup IPO* ($r = 0{,}1503$; $p < 0{,}001$) it does not suit as an instrument. As the collected variables cannot be applied as instruments and further data was not available due to limitations in access, this paper can only make theoretical assumptions about a possible instrument and consequently needs to acknowledge the endogenous nature of its main independent variable *industry relatedness* as a limitation in interpreting its results.

Wooldridge (2010) points out the suitability of exogenous factors such as time or geography as an instrument and Park and Steensma (2012) use the "availability of corporate venture capital" in a certain year to explain variations in their endogenous variable that can not be explained by the dependent variable. In the case of this paper, the exogenous factor could be the degree of formal organization and networking within a certain industry. In particular, the number of industry-specific events, congresses, trade fairs and investment pitches or the existence of dedicated lobby groups or associations could positively impact the likelihood of a CVC and a startup from the same industry to connect and to form an investor/investee relationship. On the other side, the degree of an industry's formal organization or networking probably does not impact the likelihood of a startup to go public. As this hypothesis cannot be tested in the scope of this paper, the idea and further elaborations on a suitable instrument can be subject to future research in this area.

4. Analysis

Chapter 4 explores the descriptive statistics of the dataset before presenting the results of the main probit regression and the following robustness tests.

4.1 Descriptive statistics

Table 1 explores basic summary statistics of the variables in the dataset. It shows that 11,3% out of the sample's 891 deal observations where a startup received an investment by a European CVC investor between 01.01.2002 and 31.12.2014 resulted in a *startup IPO*. Furthermore, out of the 891 deal dyads 23,6% exhibited an *industry relatedness* between the startup and the CVC investor's parent based on a 2-digit SIC code match.

Even though the sample does not solely consist of CVCs with listed parent companies, *listed CVCs* account for the majority of CVCs in the sample. In particular, 78% of deals observed include a CVC whose parent firm is publicly listed. In 11% of the deals the CVC was also acting as the *lead investor* within the investor syndicate.

The summary statistics show the heterogeneity among startup and deal characteristics, depicted by the high standard deviations among some of the variables. This observation is particularly evident in the case of the *total funding received by the startup* where the mean amount received is $75,21 Mil. with a standard deviation of 244,0 and a median of $31,24 Mil, respectively. The observed skewness supports the notion that comparably few startups that are successful (or promise to be so in the future), receive much more funding than the larger part of other startups that do not share the same success prospects. Hence, the means of this and other variables that exhibit high standard deviations, such as for the *CVC's total amount of equity invested in a* *startup*, should be interpreted cautiously. For the purpose of alleviating the skewness of these variables for the probit regression they were transformed into their logarithmic form (Table 1 Panel B).

Table 1. Sample summary statistics and variable overview

VARIABLES	**DESCRIPTION**	**(1) N**	**(2) mean**	**(3) sd**
PANEL A	VARIABLE OVERVIEW			
Dependent variable				
(1) startup_IPO	1 if startup went public	891	0.113	0.317
Main independent variable				
(2) related	1 if startup and CVC parent share the same 2-digit SIC code	891	0.236	0.425
CVC-related controls				
(3) CVC_listed	1 if CVC parent is listed at last CVC funding	891	0.780	0.414
(4) CVC_total_deals_number	Total number of deals of CVC at last CVC funding	891	72.98	65.77
Startup-related controls				
(5) startup_total_funding	Total amount of funding received by startup at last CVC funding (Mil $)	891	75.21	244.0
(6) ageatfinancing_months	Startup age at last CVC funding (in months)	891	68.89	46.38
(7) total_no_funds	Total number of funds invested in the startup at last CVC funding	891	9.544	7.079
Dyad-related controls				
(8) round_number	Last round where CVC invested in startup	891	3.756	2.782
(9) earliest_round	Earliest round where CVC invested in startup	891	2.798	2.176
(10) CVC_total_equity	Total amount CVC invested in startup (Mil $)	891	6.174	10.26
(11) CVC_lead_investor	1 if CVC investor invested in startup's first funding round and accounts for min. 50% of total equity invested	891	0.110	0.313
Year and industry controls				
(12) investment_year	Dummy variables to control for year effects at year of last investment	891	-	-
(13) 2_digit_SIC_CVC	Dummy variables to control for CVC industry fixed effects based on 2-digit SIC code classification	891	-	-
PANEL B	NATURAL LOG OF SELECTED VARIABLES			
(4) log_CVC_total_deals		891	3.729	1.225
(5) log_startup_total_funding		891	3.208	1.615
(6) log_age_at_financing		891	3.965	0.842
(7) log_total_no_funds		891	1.964	0.821
(8) log_round_number		891	1.057	0.747
(9) log_earliest_round		891	0.767	0.712
(10) log_CVC_total_equity		891	1.149	1.198

Table 2 Panel A provides an overview of the industries in which startups attracted the highest amount of funding from all external investors. The sample shows that startups from the SIC category "Business services", which includes sectors such as advertising and pre-packaged software, received the highest total amount of funding with $14.382 Mil. They were followed by startups from the "Chemicals and pharmaceuticals" industry with $12.785 Mil. and startups from "Engineering and management services" with $8.801 Mil. Panel B illustrates the CVC parent industries that invested most in startups between 2002 and 2015. Here, "Chemicals and pharmaceuticals" invested by far the highest amount in venture capital with $2.205 Mil., followed by "Communications", which comprises telecom companies ($609 Mil.), and "Printing and publishing ($460,81).

The strong concentration of CVC investments in investors from chemicals and pharmaceuticals can not just be explained by industry-specific factors such as high R&D intensity but also when taking a look on the composition of the sample. Of the 891 deal observations, 315 involve a CVC investor that belongs to the 2-digit SIC code classification 28 of chemicals and pharmaceuticals.

Besides impacting the distribution of equity invested among CVC investors, the high concentration of SIC-28 companies is further reflected in the distribution of related relationships among the industry classifications (Table 3). Of the 210 related deal relationships, 149 are composed of companies from the chemicals and pharmaceuticals sector. The consequences of this skewed distribution of related deals among SIC sectors will be further evaluated in the discussion part of this paper.

Startups from the "Chemicals and pharmaceuticals" sector account also for the largest number of startup IPOs with 40 IPOs out of 101 total IPOs in the dataset, followed by startups from "Engineering services" with 26 and "Instruments" with 11 IPOs (Table 4). Similar to the previously observed distribution of CVC equity invested in startups, the dominance of chemical companies and pharmaceuticals is even stronger on the CVC investor side. Here, the SIC-28 industry counts 73 IPOs while the two second largest investor industries, "Printing and publishing" and "Communication", only exhibit 6 IPOs each.

Table 2. Overview of the largest industries on startup- and CVC-level by investments

PANEL A: 10 largest startup industries by total funding received

Startup 2-digit SIC	**SIC description**	**Startup total funding (Mil. $)**
73	Business services including advertising and pre-packaged software	14382.02
28	Chemicals and pharmaceuticals	12785.15
87	Engineering and management services	8801.87
37	Transportation equipment	5424.17
36	Electronic and other electric equipment	5208.49
57	Furniture and home furnishing stores	4829.63
38	Instruments and related products	3757.66
59	Miscellaneous retail	1893.78
48	Communications	1850.23
30	Rubber and miscellaneous plastics	1161.76

PANEL B: 10 largest CVC industries by total equity invested in startups

CVC 2-digit SIC	**SIC description**	**Total equity invested in startup (Mil. $)**
28	Chemicals and pharmaceuticals	2205.12
48	Communications	609.06
27	Printing and publishing	460.81
35	Industrial machinery and equipment	370.99
54	Selling food for home preparation and consumption	285.39
20	Manufacturing or processing foods and beverages for human consumption	275.38
49	Electric, gas and sanitary services	261.30
36	Electronic and other electric equipment	121.34
37	Transportation equipment	116.79
13	Oil and gas extraction	112.29

Table 3. Distribution of related deal relationships among industries

Matching 2-digit SIC	Number of related deal dyad relationships
28	149
73	19
48	16
20	6
87	5
35	4
38	4
49	3
37	2
13	1
27	1
Total of 210 related CVC-startup dyad deal relationships	

Apart from the apparent dominance of chemical and pharmaceutical companies in CVC transactions, Table 2 and Table 4 offer valuable insights on how the industry landscape has changed in the first and a half decade of this century. When looking at the constituents of the startup industries in both tables we see sectors such as advertising and software (SIC 73) or engineering (SIC 26) that received the largest amount of funding and that account for the fourth and second highest number of startup IPOs, respectively. On the other side of both tables, when looking at the constituents of the CVC industries, we see telecommunications companies (SIC 48), publishers (SIC 27) or utilities (SIC 49) dominating the total equity spending of CVCs.

Table 4. Distribution of startup IPOs among startup and CVC industry classifications

Startup			CVC		
SIC 2-digit	SIC description	Number of startup IPOs	SIC 2-digit	SIC description	Number of startup IPOs
28	Chemicals and pharmaceuticals	40	28	Chemicals and pharmaceuticals	73
87	Engineering and management services	26	27	Printing and publishing	6
38	Instruments and related products	11	48	Communications	6
73	Business services including advertising and pre-packaged software	8	35	Industrial machinery and equipment	5
36	Electronic and other electric equipment	6	73	Business services including advertising and pre-packaged software	3
35	Industrial machinery and equipment	3	49	Electric, gas and sanitary services	2
57	Furniture and home furnishing stores	2	54	Selling food for home preparation and consumption	2
59	Miscellaneous retail	2	20	Manufacturing or processing foods and beverages for human consumption	1
13	Oil and gas extraction	1	36	Electronic and other electric equipment	1
48	Communications	1	50	Wholesale trade-durable goods	1
78	Motion pictures	1	80	Health services	1
Total number of startup IPOs: 101; observed time horizon between 01.01.2002 and 01.03.2020					

This observation is in line with the surge of young companies with software- and internet-based business models from the early 2000s on which attracted large sums of venture capital. Also, seeing rather established industries like telecommunications or publishing being among the largest CVC investors does support the argument for CVC as a window to new technologies and innovation for mature industries (see chapter 2.3.1).

Table 5 illustrates Pearson's correlation between all variables of the probit regression model (except year and industry dummies) including the respective p-values of the correlation coefficients. The dependent variable *startup_IPO* exhibits a low correlation (r = 0,1351) with the main independent variable *related* (p <0,001). The strongest correlation of the dependent variable that can be reported is with the natural log of the startup's total funding (r = 0,2951; p <0,001). In addition, *startup_IPO* shows significant correlations with *log_CVC_total_equity* (r= 0,2302; p <0,001) and *log_total_no_funds* (r = 0,2199; p <0,001). This underlines the view that successful startups are, first, those that receive the most funding prior an IPO, second, the ones where CVC investors are willing to invest the highest amounts of equity on an individual level, and, third, those that exhibit the largest (C)VC investor syndicates.

Unsurprisingly, the strongest overall observed correlation (r = 0,7538; p <0,001) is between the two logarithmic forms of the variables "startup total funding" and "total number of funds invested in the startup" as with more investors that invest in a company, the total amount of funding is more likely to be higher too.

Table 5. Correlations

Correlations (N = 891): correlation coefficient (p-value)											
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) startup_IPO	1.0000										
(2) related	0.1351	1.0000									
	(0.0001)										
(3) CVC_listed	0.0617	0.1098	1.0000								
	(0.0657)	(0.0010)									
(4) log_startup_total_funding	0.2951	0.0218	0.2295	1.0000							
	(0.0000)	(0.5155)	(0.0000)								
(5) log_CVC_total_equity	0.2302	0.0123	0.1124	0.6495	1.0000						
	(0.0000)	(0.7129)	(0.0008)	(0.0000)							
(6) log_round_number	0.1876	0.0268	0.2640	0.5754	0.2798	1.0000					
	(0.0000)	(0.4235)	(0.0000)	(0.0000)	(0.0000)						
(7) log_earliest_round	0.1335	-0.0194	0.2480	0.4540	0.1323	0.7975	1.0000				
	(0.0001)	(0.5631)	(0.0000)	(0.0000)	(0.0001)	(0.0000)					
(8) log_total_no_funds	0.2199	0.0558	0.2596	0.7538	0.2447	0.6682	0.5778	1.0000			
	(0.0000)	(0.0959)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)				
(9) log_age_at_financing	0.0814	0.0091	0.1797	0.1981	0.1111	0.4636	0.3553	0.2257	1.0000		
	(0.0151)	(0.7852)	(0.0000)	(0.0000)	(0.0009)	(0.0000)	(0.0000)	(0.0000)			
(10) CVC_lead_investor	-0.1144	-0.0093	-0.1943	-0.4908	-0.1453	-0.4100	-0.3787	-0.6231	-0.1040	1.0000	
	(0.0006)	(0.7821)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0019)		
(11) log_CVC_total_deals	0.1503	0.1144	0.3633	0.3288	0.2630	0.2679	0.1831	0.2921	0.0555	-0.2282	1.0000
	(0.0000)	(0.0006)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0978)	(0.0000)	

4.2 Probit regression

Before conducting a probit regression on the dichotomous variable *startup_IPO*, it is tested whether there exists a significant association between *startup_IPO* and the main independent variable *related*, which indicates an industry relatedness based on a 2-digit SIC code match. Due to the dichotomous nature of both variables the Pearson chi-squared test is being applied. The association between the variables is measured by using the two following hypotheses:

H_0: The IPO of a startup in a CVC-startup dyad relationship and the industry relatedness of the two entities are independent from each other.

H_1: The IPO of a startup in a CVC-startup dyad relationship and the industry relatedness of the two entities are not independent from each other.

Table 6. Chi-squared test

startup_IPO	related 0	1	Total
0	620	170	790
1	61	40	101
Total	681	210	891

Pearson chi2(1) = 16.2591 Pr = 0.000

According to Table 6 the likelihood chi-squared statistics is 16.2951 ($p < 0,001$). Hence, we can state at a 1% significance level that the two variables are related and reject H_0 that both are independent.

Table 7 summarizes the results of the probit regression model with the marginal effects of the coefficients. As illustrated in the table, the main independent variable and all controls were subsequently added to the model. Furthermore, the model was corrected for heteroskedasticity.

The hypothesis of this study predicted that CVC funding would be more beneficial to a startup's performance when the CVC investor's parent and the startup are from the same industry. The final regression result in Table 6 Model 12 shows an economic and statistically insignificant effect of industry relatedness (*related*) between a CVC and a target startup on the startup's IPO. Thus, the result does not support the paper's hypothesis. The pseudo R^2 of Model 12 is 0,2435. Table 6 shows that the variable *related* took an economic and statistical insignificance only after controlling for industries on the CVC side, which were added as fixed effects based on 2-digit SIC codes (Model 12). When ignoring industry fixed effects, as in Models 1 to 11, the coefficient of the variable *related* is highly statistically significant on a 1% level. If Model 11 would be interpreted, we could state on a 1% significance level that a related relationship between CVC and startup increases the likelihood of a startup IPO by 7,59%.

The SIC codes dummy variables added in Model 12 all exhibit highly statistically significant coefficients on a 1% level. As seen in Table 4, the dataset exhibits a comparably small number of SIC industries on the CVC side that account for the observed startup IPOs in this sample. Hence, there are CVC SIC industry groups that consist exclusively of deals where no startup IPO was observed. As the outcome of the dependent variable will be always zero for these observations, they don't possess any explanatory power for the likelihood of the occurrence of an IPO and were therefore automatically dropped from the sample, resulting in the 758 observations in Model 12.

The variable *log startup total funding* positively impacts the likelihood of a startup IPO and is highly statistically significant at a 1% level throughout all models. A one million increase in the funding of a startup increases the likelihood of a startup IPO by 5%. The *log of total equity invested* of a CVC in a particular startup shows a weak positive economic effect at a 10% significance level in Models 4 to 10. When adding year and industry controls, however, this variable is rendered statistically insignificant. *Startup age* at the last financing round has also a weak positive effect on the startup IPO in Model 11 and 12 which is significant at a 10% and at a 5% level, respectively. The other control variables of the probit analysis are statistically insignificant at every stage of the model. Thus, an influence of these variables on the likelihood of a startup IPO cannot be inferred.

Table 7. Probit regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
related	0.101***	0.0950***	0.0766***	0.0745***	0.0743***	0.0742***	0.0733***	0.0731***	0.0723***	0.0688***	0.0759***	0.00908
	(0.0292)	(0.0287)	(0.0255)	(0.0256)	(0.0256)	(0.0258)	(0.0258)	(0.0255)	(0.0256)	(0.0256)	(0.0262)	(0.0216)
CVC_listed		0.0359	-0.00409	-0.00202	-0.00627	-0.00614	-0.00736	-0.00968	-0.0101	-0.0224	-0.0187	-0.0583
		(0.0235)	(0.0237)	(0.0237)	(0.0252)	(0.0253)	(0.0253)	(0.0255)	(0.0251)	(0.0273)	(0.0258)	(0.0442)
log_startup_total_funding [1]			0.0564***	0.0488***	0.0460***	0.0461***	0.0429***	0.0443***	0.0435***	0.0426***	0.0411***	0.0500***
			(0.00646)	(0.00731)	(0.00840)	(0.00846)	(0.0107)	(0.0106)	(0.0104)	(0.0101)	(0.00956)	(0.0117)
log_CVC_total_equity [1]				0.0158*	0.0162*	0.0161*	0.0176*	0.0176*	0.0177*	0.0166*	0.0132	0.0108
				(0.00836)	(0.00835)	(0.00865)	(0.00935)	(0.00928)	(0.00920)	(0.00910)	(0.00912)	(0.00997)
log_round_number					0.0106	0.0116	0.00891	-0.00214	-0.00215	-0.00619	-0.0112	-0.00815
					(0.0133)	(0.0193)	(0.0199)	(0.0211)	(0.0209)	(0.0211)	(0.0204)	(0.0228)
log_earliest_round						-0.00144	-0.00198	-0.00182	-0.00201	-0.000447	0.00411	0.00992
						(0.0183)	(0.0183)	(0.0181)	(0.0179)	(0.0179)	(0.0177)	(0.0191)
log_total_no_funds [1]							0.00884	0.00924	0.00626	0.00629	0.00564	-0.0115
							(0.0182)	(0.0180)	(0.0192)	(0.0190)	(0.0184)	(0.0217)
log_age_at_financing [1]								0.0162	0.0166	0.0183	0.0215*	0.0254**
								(0.0113)	(0.0113)	(0.0113)	(0.0111)	(0.0125)
CVC_lead_investor									-0.0241	-0.0201	-0.0191	-0.0442
									(0.0482)	(0.0498)	(0.0454)	(0.0301)
log_CVC_total_deals [1]										0.0111	0.00842	0.00105
										(0.00786)	(0.00745)	(0.0100)
_Iinvestmen_2003											-0.0173	-0.00885
											(0.0529)	(0.0602)
_Iinvestmen_2004											0.0661	0.131
											(0.0871)	(0.123)
_Iinvestmen_2005											0.0897	0.174
											(0.102)	(0.148)
_Iinvestmen_2006											0.0996	0.136
											(0.100)	(0.122)
_Iinvestmen_2007											0.0116	0.0220
											(0.0613)	(0.0707)
_Iinvestmen_2008											-0.0246	-0.00167
											(0.0436)	(0.0633)
_Iinvestmen_2009											0.0266	0.0475
											(0.0680)	(0.0848)
_Iinvestmen_2010											0.0189	0.0164
											(0.0748)	(0.0732)
_Iinvestmen_2011											0.0475	0.0524
											(0.0766)	(0.0855)
_Iinvestmen_2012											0.0170	0.0353
											(0.0620)	(0.0778)
_Iinvestmen_2013											0.0984	0.0931
											(0.0895)	(0.0942)
_Iinvestmen_2014											0.0448	0.0393
											(0.0635)	(0.0661)
_I_digit_CV_5												0.950***
												(0.0159)
_I_digit_CV_6												0.968***
												(0.0197)
_I_digit_CV_7												0.933***
												(0.0643)
_I_digit_CV_10												0.968***
												(0.0256)
_I_digit_CV_11												0.944***
												(0.0170)
_I_digit_CV_16												0.965***
												(0.0455)
_I_digit_CV_17												0.949***
												(0.0135)
_I_digit_CV_18												0.939***
												(0.0117)
_I_digit_CV_19												0.937***
												(0.0197)
_I_digit_CV_24												0.960***
												(0.00875)
_I_digit_CV_26												0.937***
												(0.0120)
Observations	891	891	891	891	891	891	891	891	891	891	891	758

* significance at the 10% level; ** significance at the 5% level; ***significance at the 1% level. [1] at time of last CVC investment in the focal startup. Pseudo R^2: 0,2435

4.3 Robustness tests

Several tests were conducted to ensure the robustness of the results achieved in the probit regression. While the first and the second robustness tests used smaller and larger sample sizes, respectively, the third and fourth used the initial sample but different measures of calculating the main independent variable *industry relatedness* by using a 4-digit SIC code match and the Fama French 12 industry classification, respectively.

The first robustness check takes a subsample of the dataset used in the main regression of this paper. The subsample consists solely of observations where the parent company of the CVC was publicly listed at the time of its last investment in the startup. The original dataset included deal observations with privately-owned CVC parent companies which often do not disclose financial information like total assets. Hence, in the initial regression it was not possible to control for the CVC parent's size by using total assets or market capitalization as done usually in CVC research. The subsample has 694 observations. The control variable *CVC total assets*, which is the total assets of a CVC parent company in thousand $ at the year prior to the last investment in the associated startup, is added with the natural log to the model. The values for total assets were retrieved from Datastream in the CVC parents' local currencies and subsequently converted to USD taking the exchange rate of 31.12.XX of the respective year via conversion tables provided on ofx.com. The binary control variable *CVC listed* is consequently removed from the model. Again, Pearson's chi-squared test is conducted with a statistical significance level at 1%, thus rejecting the null hypothesis that *startup IPO* and *related* are independent from each other. As in the main regression, all variables were added subsequently to the model and the model was corrected for heteroskedasticity. The result of the probit regression (Appendix 2) exhibits a statistically insignificant coefficient for *related* and therefore supports the results of the main regression. However, unlike in the main regression where the coefficient was highly statistically significant until adding industry dummies, *related* is only weakly statistically significant at a 10% level before adding industry dummies (Model 11). The newly added control for a *CVC parent's total assets* has a negative impact ($ß = -0,028$; $p <0,05$) on the likelihood of the target startup's IPO.

The second robustness test was a probit regression conducted on a larger dataset of European CVC investment deals. Due to missing data in Thomson OneBanker, this robustness test does not control for *startup age* at financing round, the *CVC's total equity* invested in a startup as well as not for the binary variable *CVC lead investor*. As for the main regression, observations of non-listed CVC parents did not allow to control for *CVC total assets* as a proxy for a CVC parent's size. The Pearson chi-squared test for *startup IPO* and *related* was significant at a 1% level. However, the marginal effect of the probit regression (Appendix 3) has a statistically insignificant coefficient, supporting the result of the main regression.

A third robustness test used a 4-digit SIC code match of the CVC and the startup to compute the main independent variable *related*. The results are 67 deal observations that exhibited an industry relatedness (compared to 210 when using 2-digit SIC codes). A Pearson chi-squared test indicated a significant association of *startup IPO* and *related* on a 1% level. Aside from controlling for industry on a 4-digit SIC code level, the same control variables were used as in the main regression. Again, the results of the probit model (Appendix 4) showed no effect of industry relatedness in a CVC-startup dyad on the likelihood of a startup IPO.

Finally, a last robustness test was conducted on the initial sample but using the Fama French 12 industry classification (FF12) to calculate the variable *related*. FF12 enables a more concentrated classification of industries and is therefore of benefit when analyzing smaller samples (Bhojraj et al., 2003). It classifies 11 industry categories and one "other" category on a scale from 1 to 12. A conversion table (Appendix 5) was used to translate the SIC codes of the sample to FF12 codes. After conducting Pearson's chi-squared test (p <0,001), the same regression model was used as in the main regression, only substituting the SIC code-based industry control dummies by FF12 control dummy variables. The different industry classification led to 236 observed deal dyads that exhibited an industry relatedness (compared to 210 when using 2-digit SIC code matches). In line with the previous results, the impact of industry relatedness on the likelihood of a startup IPO was statistically insignificant (Appendix 6).

5. Discussion

This study examined the effect of industry relatedness between a CVC parent and a target startup on the subsequent performance of the startup. To define industry relatedness, it used a 2-digit SIC code-based matching of the two entities to calculate the dichotomous variable *related*. As for the dependent variable, it followed previous researchers and used the IPO of the startup as a proxy for startup and deal performance. The model analyzed a dataset of 891 startup deals of European headquartered CVCs, retrieved from the Thomson OneBanker database. The results of this paper show no significant effect of industry relatedness between a CVC and an invested startup on the startup's likelihood to perform an IPO. In particular, the *related* coefficients of the marginal effects were rendered insignificant after adding industry control variables which were in turn highly statistically significant. This could be partially explained by the observed skewed distribution of SIC industries in terms of related deal pairs and, foremost, in terms of startup IPOs. As already seen in the descriptive statistics section, certain industries not only account for a large part of the CVC investor body in this sample but they also exhibit the majority share of startup IPOs. Hence, within this sample, the predominant drivers for a startup IPO were the industry of the CVC investors alongside the total funding of a startup. Apparently, industries that engage more often in CVC such as pharmaceuticals and chemicals are also more successful when it comes to achieving IPOs of the startups they invested in. While other factors might as well play a role here, it confirms the notion that investors that engage more frequently in CVC tend to be more successful in doing so (Benson & Ziedonis, 2009). Also, it is in line with previous findings such as Basu et al. (2011) that industries with rapid technological change and a high competitive intensity engage more in CVC activities like pharmaceuticals that rely heavily on new drug development or telecoms that stand in fierce competition with their peers for new subscribers. However, the dominance of relatively few industries in explaining the outcome variable of this analysis limits the interpretation of the results and will be further assessed in the course of this chapter. Subsequently performed robustness tests also exhibited insignificant results and confirmed the results of the main analysis.

The study contributes to the existent literature mainly in two ways. First, it contributes to the explanation of potential drivers for CVC investment success by adding the concept of industry relatedness. While previous studies used industry relatedness to explain decision making of CVCs and startups ex-ante the CVC investment, this paper, to the best of the author's knowledge, is first to analyze on how industry relatedness can influence CVC deal outcomes ex post once a CVC and a startup entered a deal relationship. Secondly, it adds a European perspective to CVC literature. The majority of research conducted in this domain is based upon US datasets. European CVC activities, however, are comparably less analyzed despite the continent's strong activities in CVC, especially in industries such as pharmaceuticals, telecommunications or media. Hence, it can help providing a more balanced view on CVC activities on a global level as well as identifiying possible differences in CVC characteristics between the US and Europe.

This thesis is subject to several limitations that need to be acknowledged. These can be grouped into limitations related to the dataset and methodology-related limitations. The dataset-related limitations concern the database Thomson OneBanker. As already mentioned previously and as remarked by several scholars before, Thomson OneBanker's database of (C)VC activities contains inconsistencies, missing data for certain categories and is not exhaustive of all CVC deals. Also, sum of the figures like investment sums or total funding of startups are subject to estimations. Datasets of European deals are particularly affected by this caveat, compared to the more exhaustive US data. While this study's access to data was limited to the Thomson OneBanker database, other CVC scholars further rely on complementary databases such as Lexus/Nexus, Zephyr, VICO or proprietary datasets to alleviate this issue. Another dataset-related issue is the strong concentration of certain industries like pharmaceuticals in the sample. While this can certainly be explained by the fact that these industries are particularly suited for and attracted by venture capital investments, it might also be that other industries are underrepresented in the dataset due to missing data. A case of non-randomly missing data, which would bias the results of the analysis, could not be ruled out entirely due to the private nature of CVC and limitations to access alternative data sources.

Among the methodology-related caveats of this study, the first one to mention is the SIC code-based industry relatedness calculation. The shortcomings of this method were already elaborated in more detail in chapter 3.2.3. Here, the classification of companies with diverse products and services under one SIC code is certainly one of the main caveats of the SIC code-based approach, as well as the question on whether the SIC methodology is able to reflect the nature of businesses that evolved in the 21st century such as internet- or data-related business models. As this paper analyzed European CVCs, other classification registers like the CorpTech database, which only displays US businesses, could not be used.

The second methodological limitation is the endogenous nature of the main independent variable industry relatedness. As illustrated in chapter 3.2.4, there are means to dissolve this issue like the application of a bivariate probit model with an instrumental variable regression as done by Park and Steensma (2012). Due to limited data available, this procedure could not be followed within the scope of this work and only exercised on a theoretical level. Related to this issue, Masulis and Nahata (2009) report that the decision of a CVC to invest in a startup and the type of financial contract they choose could be a result of the startup performance prior the investment. This ex post selection bias could imply that CVCs would only invest in startups that already perform well. This should be considered when measuring the impact of the nature of a CVC-startup dyad on the success of the CVC investment as it is done in this study. However, a further robustness test was not within the scope of this work considering the limited data on Thomson OneBanker for a startup's pre-IPO performance and further resource restrictions. Lastly, the author acknowledges that the results of this work could be subject to omitted variable bias, in particular, the missing of other determinant variables of a startup IPO that could be correlated with industry relatedness.

The results and the nature of this study leave manifold potentials for future researchers. First, the concept of industry relatedness could be applied to performance measures in CVC other than the startup IPO, exploring possible significant impacts. Also, the use of more exhaustive CVC datasets could be a means to test the results of this paper. Here, time can play an important factor. The analyzed timeframe of this paper was 2002-2015, a time where the CVC industry was mainly dominated by investors from the pharmaceuticals, chemicals and telecoms sectors. Meanwhile, more industries engage regularly in CVC and have identified CVC as a strategic-relevant activity to foresee possible disruptions as the examples of H&M and Ikea show (Financial Times, 2019). Future research could analyze this more diverse CVC investor body and the role of industry relatedness on the performance of its investments.

Methodology-wise, future research can use other means to determine industry relatedness. These could include qualitative measurements like in-person interviews, questionnaires or the screening of the CVC parent's and startup's websites, IPO prospectus or annual reports. Also, more sophisticated models like the survivor principle-based approach proposed by Cefis and Rigamonti (2018) could be applied. The research methodology of this paper could be also applied to US data which, in general, is more exhaustive compared to European CVC deal data. Besides the higher availability of data, it would further allow to use industry classifications that are deemed more precise than SIC codes but that only exist for US markets, such as the CorpTech Directory.

6. Conclusion

This study could not empirically prove a significant relationship between industry relatedness of a CVC parent and a startup and the subsequent startup performance measured by the occurrence of an IPO. The results were generated by using a probit model and were further confirmed by several robustness tests. The sample of European CVC deals between 2002 and 2015 exhibited a strong concentration of CVC investors from few industries like chemicals and pharmaceuticals. Meanwhile, the European CVC industry has more diversified and the increasing sums of CVC investments underline the strategic importance of CVC for traditional and new players in this field. In the years to come, future research could harness new data based on a more heterogenous CVC landscape. Together with the positive impact of related investments that is largely reported in the M&A literature, this should encourage future researchers to apply this concept to analyze the performance drivers of CVC investments.

Bibliography

Achleitner, A.-K., Braun, R. and Lutz, E. (2014). Industry relatedness in trade sales and venture capital investment returns. *Small Business Economics,* 43: 621-637

Anand, B. and Galetovic, A. (2004). Strategies that work when property rights don't. *Advances in the Study of Entrepreneurship, Innovation, and Economic Growth*, Vol. 15, Gary Libecap (Ed)

Basu, S., Phelps, C. and Kotha, S. (2011). Towards understanding who makes corporate venture capital investments and why. *Journal of Business Venturing,* 26(2): 153-171

Belderbos, R., Jacob, J. and Lokshin, B. (2018). Corporate venture capital (CVC) investments and technological performance: Geographic diversity and the interplay with technology alliances. *Journal of Business Venturing*, 33(1): 20-34

Benson, D. and Ziedonis, R. H. (2009). Corporate venture capital as a window on new technologies: Implications for the performance of corporate investors when acquiring startups. *Organization Science*, 20(2): 329-351

Bhojraj, S., Lee, C. and Oler, D. (2003). What's my line? A comparison of industry classification schemes for capital market research. *Journal of Accounting Research*, 41(5): 745-774

Birkinshaw, J. and Hill, S.A. (2005). Corporate venturing units: Vehicles for success in the new Europe. *Organizational Dynamics*, 34(3): 247-257

Capron, L. (1999). The long-term performance of horizontal acquisitions. *Strategic Management Journal*, 20: 987-1018

Cassiman, B., Colombo, M., Garrone, P. and Veugelers, R. (2003). The impact of M&A on the R&D process. An empirical analysis of the role of technological and market relatedness. *Katholieke Universiteit Leuven Research Paper*

CB Insights (2019). The Most Active Corporate VC Firms Globally. Accessed 06 May 2020, https://www.cbinsights.com/research/corporate-venture-capital-active-2014/

Cefis, E. and Rigamonti, D. (2018). A unidirectional measure of industry relatedness and its application to acquisitions framework. *International Journal of Management Sciences and Business Research*, 7(9): 2226-2235

Chang, S., Eggers, J.P. and Keum, D.D. (2017). Market relatedness, technological relatedness, and the dynamics of organizational expansion. *Colombia Business School Research Paper,* No.14-60

Chemmanur, T. J., Loutskina, E. and Tian, X. (2012). Differences in financial market responses to backing by corporate versus independent venture capitalists. Working paper, Boston College

Chemmanur, T. J., Loutskina, E. and Tian, X. (2014). Corporate venture capital, value creation, and innovation. *The Review of Financial Studies,* 27 (8): 2434-2473

Chesbrough, H. (2002). Making sense of corporate venture capital. *Harvard Business Review,* March 2002 issue

Cochrane, J. H. (2005). The risk and return of venture capital. *Journal of Financial Economics*, 75: 3-52

Cohen, W. M. and Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35: 128-152

Colombo, M. G. & Murtinu, S. (2017). Venture capital investments in Europe and portfolio firm's economic performance: Independent versus corporate investors. *Journal of Economics & Management Strategy,* 26 (1): 35-66

Dauderstaedt, P. (2013). Success factors in strategic corporate venturing.

Dimitrova, L. (2015). *Strategic acquisitions by corporate venture capital investors.* Available at SSRN: https://ssrn.com/abstract=2553786

Dushnitsky, G. and Lenox, M. J. (2005). When do incumbents learn from entrepreneurial ventures? Corporate venture capital and investing firm innovation rates. *Research Policy,* 34: 615-639

Dushnitsky, G. and Lenox, M. J. (2006). When does corporate venture capital investment create firm value? *Journal of Business Venturing,* 21(6): 753-772

Dushnitsky, G. and Shaver, J. M. (2005). Limitations to inter-organizational knowledge acquisition: the paradox of corporate venture capital. *Strategic Management Journal*, 30(10): 1045-1064

Financial Times (2019). Ikea and H&M step up venture capital investments. Accessed 08 May 2020, https://www.ft.com/content/aa8d04f4-fca5-11e9-a354-36acbbb0d9b6

Fleßner, P., Wade, M. and Obewegeser, N. (2019). Making corporate venture capital work. *MIT Sloan Management Review,* 06/2018

French, K. (2020). Detail for 12 industry portfolios. Accessed 04 May 2020, https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_12_ind_port.html

Gans, J. S., Hsu, D. and Stern, S. (2001). When does startup innovation spur the gale of creative destruction? *Rand Journal of Economics,* 33: 571-586

Gompers, P. and Lerner, J. (2000a). The determinants of corporate venture capital success: organizational structure, incentives and complementarities. *Concentrated Corporate Ownership,* University of Chicago Press for the National Bureau of Economic Research: 17-50

Gompers, P. and Lerner, J. (2000b). Money chasing deals? The impact of fund inflows on private equity valuations. *Journal of Financial Economics*, 55: 281-325

Gondhalekar, V. B., Raymond Sant, R. and S.P. Ferris (2004). The price of corporate acquisition: Determinants of cash takeover premia. *Applied Economics Letters*, 11(12): 735-739

Guo, B., Lou, Y. and Pérez-Castrillo, D. (2015). Investment, duration, and exit strategies for corporate and independent venture capital-backed start-ups. *Journal of Economics & Management Strategy,* 24(2): 415-455

Ivanov, V. I. and Xie, F. (2010). Do corporate venture capitalists add value to start-up firms? Evidence from IPOs and acquisitions of VC-backed companies. *Financial Management*, 39(1): 129-152

Kaplan, S. N. and Strömberg, P. (2003). Financial contracting theory meets the real world: an empirical analysis of venture capital contracts. *Review of Economic Studies*, 78: 281-315

Koehn, A. (2018). Where entrepreneurship and finance meet: Startup valuation and acquisition in the venture capital and corporate context

Lahr, H., and Mina, A. (2016). Venture capital investments and the technological performance of portfolio firms. *Research Policy*, 45: 303-318

Lee, U.S., Park, G. and Kang, J. (2018). The double-edged effects of the corporate venture capital unit's structural autonomy on corporate investor's explorative and exploitative innovation. *Journal of Business Research,* 88: 141-149

Lien, L. B. and Klein, P. G. (2009). Using competition to measure relatedness. *Journal of Management*, 35(4): 1078-1107

Lien, L. B. and Klein, P. G. (2013). Can the survivor principle survive diversification? *Organization Science*, 24(5): 1478-1494

Lim, M.-H. and Lee, J.-H. (2016). The effects of industry relatedness and takeover motives on cross-border acquisition completion. *Journal of Business Research,* 69: 4748-4792

MacMillan, I., Roberts, E., Livada, V., and Wang, A. (2008). Corporate Venture Capital (CVC): Seeking innovation and strategic growth. National Institute of Science and Technology report (NIST GCR 08-916)

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71-87

Masulis, R. W. and Nahata, R. (2009). Financial contracting with strategic investors: Evidence from corporate venture capital backed IPOs. *Journal of Financial Intermediation*, 18 (4): 599-631

Maula, M. V. J., (2001). Corporate venture capital and the value-added for technology-based new firms. 951-22-5780-7

Maula, M. V. J. and Murray, G. (2001). Complementary value-adding roles of corporate venture capital and independent venture capital investors. *Working Paper: Helsinki University of Technology, Institute of Strategy and International Business*

Metrick, A. and Yasuda, A. (2011). Venture capital & the finance of innovation. *John Wiley & Sons, Inc.*, 2nd edition, 3-21

Nerkar, A. and Roberts, P. W. (2004). Technological and product-market experience and the success of new product introductions in the pharmaceutical industry. *Strategic Management Journal,* 25 (8-9): 779-799

Oxley, J. and Sampson, R. (2004). The scope and governance of international R&D alliances. *Strategic Management Journal,* Special Issue 25 (8-9): 723-749

Park, H. D. and Steensma, H. K. (2012). When does corporate venture capital add value for new ventures? *Strategic Management Journal*, 33: 1-22

Schefczyk, M. and Gerpott, T. J. (2002). Management support for portfolio companies of venture capital firms: An empirical study of German venture capital investments. *British Journal of Management,* 201-216

Siren, C. A., Kohtamäki, M. and Kuckertz, A. (2012). Exploration and exploitation strategies, profit performance, and the mediating role of strategic learning: Escaping the exploitation trap. *Strategic Entrepreneurship Journal,* 6(1): 18-41

Sosa, M. L. (2009). Application-specific R&D capabilities and the advantage of incumbents. Evidence from the anticancer drug market. *Management Science,* 55(8): 1409-1422

Stellner, F. (2015). The impact of technological distance on M&A target choice and transaction value. *Max Planck Institute for Innovation and Competition Research Paper,* No. 15-12

Teece, D.J., Rumelt, R., Dosi, G. and Winter, S. (1994). Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior & Organization*, 23(1): 1-30

UK Office for National Statistics (2007). UK SIC 2007. Accessed 26 February 2020, https://www.ons.gov.uk/methodology/classificationsandstandards/ukstandardindustrialclassificationo ukstandardindustria/uksic2007

Valentini, G. and Di Guardo, M. C. (2012). M&A and the profile of inventive activity. *Strategic Organization*, 10(4), 384-405

Villalonga, B. and McGahan, A. M. (2005). The choice among acquisitions, alliances and divestitures. *Strategic Management Journal,* 26: 1183-1208

Wadhwa, A., Phelps, C. and Kotha, S. (2015). Corporate venture capital portfolios and firm innovation. *Journal of Business Venturing*, 13(1): 95-112

Waguespack, D. and Fleming, L. (2009). Scanning the commons? Evidence on the benefits to startups participating in open standards development. *Management Science*, 55(2): 210-223

Weber, C., Bauke, B. and Raibulet, V. (2016). An empirical test of the relational view in the context of corporate venture capital: relational view and corporate venture capital. *Strategic Entrepreneurship Journal*, 10(3):274-299

Wong, P. and O’Sullivan, N. (2001). The determinants and consequences of abandoned takeovers. *Journal of Economic Surveys,* 15(2): 145-186

Wooldridge, J. M. (2010). *Econometric Analysis of cross section and panel data.* MIT Press: Cambridge, MA, 2nd Edition

Yang, Y., Chen, T. & Zhang, L. (2016). Corporate venture capital program autonomy, corporate investor’s attention and portfolio diversification. *Journal of Strategy and Management,* 9(3): 302

Appendix

Appendix 1 - Main regression (Model 12)

```
Probit regression, reporting marginal effects           Number of obs =    758
                                                        Wald chi2(32) =      .
                                                        Prob > chi2   =      .
Log pseudolikelihood = -225.06286                       Pseudo R2     = 0.2435
```

startu~O	dF/dx	Robust Std. Err.	z	P>|z|	x-bar	[95% C.I.	]
related*	.0090754	.021633	0.43	0.668	.261214	-.033324	.051475
CVC_li~d*	-.0582521	.0442428	-1.58	0.113	.825858	-.144966	.028462
log_st~g	.0499922	.0117082	4.05	0.000	3.31259	.027045	.07294
log_CV~y	.010796	.0099729	1.08	0.281	1.21609	-.00875	.030342
log_ro~r	-.0081546	.0227531	-0.36	0.722	1.08397	-.05275	.036441
log_ea~d	.0099197	.0190834	0.52	0.604	.782279	-.027483	.047323
log_to~s	-.0114566	.0216531	-0.53	0.596	2.00798	-.053896	.030983
log_ag~g	.0254098	.012493	2.07	0.038	3.95153	.000924	.049896
CVC_le~r*	-.044151	.0301133	-1.01	0.315	.100264	-.103172	.01487
log_CV~s	.0010477	.0100328	0.10	0.917	3.87938	-.018616	.020712
_Ii~2003*	-.0088507	.0601657	-0.14	0.889	.043536	-.126773	.109072
_Ii~2004*	.1308815	.1227651	1.47	0.141	.056728	-.109734	.371497
_Ii~2005*	.1742088	.1479617	1.69	0.092	.042216	-.115791	.464208
_Ii~2006*	.1361437	.121974	1.54	0.124	.055409	-.102921	.375208
_Ii~2007*	.0220319	.0706668	0.34	0.732	.077836	-.116473	.160536
_Ii~2008*	-.001671	.0633068	-0.03	0.979	.077836	-.12575	.122408
_Ii~2009*	.0474742	.0848246	0.67	0.503	.05409	-.118779	.213727
_Ii~2010*	.0164339	.0732049	0.24	0.807	.068602	-.127045	.159913
_Ii~2011*	.0524358	.0854596	0.74	0.461	.080475	-.115062	.219934
_Ii~2012*	.0353053	.0777554	0.52	0.601	.081794	-.117092	.187703
_Ii~2013*	.0930829	.0942295	1.28	0.202	.110818	-.091603	.277769
_Ii~2014*	.0393067	.0661205	0.67	0.502	.199208	-.090287	.168901
_I_di~_5*	.9501378	.0159126	4.72	0.000	.038259	.91895	.981326
_I_di~_6*	.967708	.0197436	5.17	0.000	.097625	.929011	1.0064
_I_di~_7*	.9333575	.064267	6.21	0.000	.415567	.807396	1.05932
_I_di~10*	.9681129	.0255601	5.03	0.000	.114776	.918016	1.01821
_I_di~11*	.9439226	.0170273	4.59	0.000	.026385	.91055	.977295
_I_di~16*	.9652146	.0455042	5.23	0.000	.211082	.876028	1.0544
_I_di~17*	.9486272	.0134908	4.90	0.000	.031662	.922186	.975069
_I_di~18*	.9392081	.0117076	.	.	.005277	.916262	.962155
_I_di~19*	.9371566	.0197282	4.28	0.000	.015831	.89849	.975823
_I_di~24*	.9595751	.0087498	6.21	0.000	.039578	.942426	.976724
_I_di~26*	.9373292	.0120148	5.14	0.000	.002639	.913781	.960878
obs. P	.1332454						
pred. P	.0644574	(at x-bar)					

```
(*) dF/dx is for discrete change of dummy variable from 0 to 1
    z and P>|z| correspond to the test of the underlying coefficient being 0
```

Appendix 2 – Robustness test with subsample of listed CVCs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
related	0.116***	0.105***	0.0607**	0.0549**	0.0543**	0.0529**	0.0509**	0.0516**	0.0581**	0.0482*	0.0549*	0.00877
	(0.0329)	(0.0327)	(0.0269)	(0.0259)	(0.0259)	(0.0262)	(0.0260)	(0.0258)	(0.0292)	(0.0282)	(0.0289)	(0.0288)
log_CVC_total_		-0.0188**	-	-0.0243***	-	-	-	-	-0.0311***	-	-0.0294***	-
assets			0.0245***		0.0247***	0.0254***	0.0254***	0.0258***		0.0324***		0.0280*
												*
		(0.00893)	(0.00796)	(0.00781)	(0.00780)	(0.00788)	(0.00789)	(0.00775)	(0.00954)	(0.00954)	(0.00968)	(0.0136)
log_startup_total			0.0688***	0.0570***	0.0535***	0.0540***	0.0458***	0.0470***	0.0544***	0.0510***	0.0452***	0.0560*
_funding												**
			(0.00859)	(0.00919)	(0.0101)	(0.0102)	(0.0123)	(0.0122)	(0.0144)	(0.0138)	(0.0129)	(0.0181)
log_CVC_total_				0.0263***	0.0265***	0.0253**	0.0289***	0.0287***	0.0357***	0.0331***	0.0289**	0.0250*
equity												
				(0.00984)	(0.00977)	(0.0100)	(0.0103)	(0.0102)	(0.0120)	(0.0117)	(0.0116)	(0.0138)
log_round_numb					0.0122	0.0204	0.0138	-0.000932	0.000767	-0.00702	-0.0136	-0.0235
er												
					(0.0148)	(0.0227)	(0.0226)	(0.0249)	(0.0300)	(0.0298)	(0.0294)	(0.0343)
log_earliest_roun						-0.0109	-0.0120	-0.0119	-0.0169	-0.0131	-0.00470	0.00639
d												
						(0.0215)	(0.0215)	(0.0213)	(0.0251)	(0.0250)	(0.0249)	(0.0292)
log_total_no_fun							0.0220	0.0239	0.0186	0.0176	0.0218	0.00662
ds												
							(0.0203)	(0.0199)	(0.0252)	(0.0249)	(0.0240)	(0.0317)
log_age_at_finan								0.0209	0.0256	0.0275	0.0269	0.0348*
cing												
								(0.0143)	(0.0179)	(0.0175)	(0.0168)	(0.0185)
log_CVC_total_										0.0252**	0.0221*	0.00074
deals												9
										(0.0117)	(0.0113)	(0.0137)
_Iinvestmen_200											0.0425	0.0373
3												
											(0.115)	(0.122)
_Iinvestmen_200											0.156	0.141
4												
											(0.151)	(0.158)
_Iinvestmen_200											0.247	0.241
5												
											(0.187)	(0.202)
_Iinvestmen_200											0.163	0.137
6												
											(0.153)	(0.158)
_Iinvestmen_200											0.0862	0.0457
7												
											(0.123)	(0.115)
_Iinvestmen_200											0.0558	0.0231
8												
											(0.117)	(0.108)
_Iinvestmen_200											0.127	0.0790
9												
											(0.141)	(0.132)
_Iinvestmen_201											0.0608	0.00415
0												
											(0.129)	(0.103)
_Iinvestmen_201											0.149	0.102
1												
											(0.147)	(0.143)
_Iinvestmen_201											0.0984	0.0523
2												
											(0.131)	(0.123)
_Iinvestmen_201											0.206	0.145
3												
											(0.152)	(0.150)
_Iinvestmen_201											0.126	0.0747
4												
											(0.114)	(0.110)
_I_digit_CV_3												0.00118
												(0.102)
_I_digit_CV_4												-0.0250
												(0.0925)
_I_digit_CV_5												0.127
												(0.0851)
_I_digit_CV_7												0.0237

												(0.100)
_I_digit_CV_12												0.0252
												(0.0888)
_I_digit_CV_14												0.0163
												(0.146)
Observations	694	694	694	694	694	694	694	694	640	640	640	565

Appendix 3 – Robustness test with larger sample size

```
Probit regression, reporting marginal effects           Number of obs =    919
                                                        Wald chi2(30) =      .
                                                        Prob > chi2   =      .
Log pseudolikelihood = -263.09675                       Pseudo R2     = 0.2185
```

Startu~O	dF/dx	Robust Std. Err.	z	P>|z|	x-bar	[95% C.I.	]
related*	.0061849	.0204428	0.31	0.758	.250272	-.033882	.046252
Listed~C*	-.029487	.0323169	-1.00	0.318	.803047	-.092827	.033853
log_st~g	.0437366	.0074361	5.68	0.000	3.10267	.029162	.058311
log_ro~r	.0158291	.0193284	0.82	0.410	1.03551	-.022054	.053712
log_ea~d	.016802	.0181122	0.93	0.352	.768383	-.018697	.052301
log_to~s	-.0076262	.0163715	-0.47	0.639	1.94045	-.039714	.024461
log_CV~s	.0020078	.0088549	0.23	0.821	3.76976	-.015347	.019363
_Ii~2003*	-.0369365	.0329644	-0.83	0.404	.04679	-.101546	.027673
_Ii~2004*	.0243974	.0589481	0.46	0.645	.059848	-.091139	.139933
_Ii~2005*	.0815525	.0878308	1.20	0.231	.043526	-.090593	.253698
_Ii~2006*	.0727602	.0774619	1.18	0.237	.052231	-.079062	.224583
_Ii~2007*	-.0165444	.0389959	-0.39	0.699	.075082	-.092975	.059886
_Ii~2008*	-.0221316	.0389297	-0.49	0.622	.07617	-.098432	.054169
_Ii~2009*	.0231085	.0598904	0.43	0.669	.050054	-.094275	.140492
_Ii~2010*	.0117316	.056102	0.22	0.825	.068553	-.098226	.121689
_Ii~2011*	.0081893	.0490307	0.17	0.862	.08161	-.087909	.104288
_Ii~2012*	-.012852	.0411615	-0.29	0.771	.08161	-.093527	.067823
_Ii~2013*	.0385342	.053943	0.82	0.415	.119695	-.067192	.14426
_Ii~2014*	.0009014	.0399513	0.02	0.982	.18716	-.077402	.079204
_I_di~_3*	.9432556	.0121662	4.69	0.000	.021763	.91941	.967101
_I_di~_6*	.9472195	.0145789	5.01	0.000	.033732	.918646	.975794
_I_di~_7*	.970878	.0145549	5.75	0.000	.100109	.942351	.999405
_I_di~_8*	.9513826	.0484988	6.70	0.000	.379761	.856327	1.04644
_I_di~11*	.9669571	.0217456	5.49	0.000	.106638	.924337	1.00958
_I_di~12*	.9413138	.0162869	4.89	0.000	.023939	.909392	.973236
_I_di~17*	.9618107	.0450938	5.79	0.000	.228509	.873429	1.05019
_I_di~18*	.9462375	.0108286	5.44	0.000	.026115	.925014	.967461
_I_di~19*	.9382539	.0096959	.	.	.005441	.91925	.957258
_I_di~20*	.9394746	.0135088	4.82	0.000	.016322	.912998	.965951
_I_di~27*	.9637849	.0072705	6.26	0.000	.054407	.949535	.978035
_I_di~29*	.9362299	.0099703	5.55	0.000	.002176	.916688	.955771
obs. P	.1196953						
pred. P	.0653543	(at x-bar)					

```
(*) dF/dx is for discrete change of dummy variable from 0 to 1
    z and P>|z| correspond to the test of the underlying coefficient being 0
```

Appendix 4 – Robustness test with 4-digit SIC code-based relatedness calculation

```
Probit regression, reporting marginal effects           Number of obs =    758
                                                        Wald chi2(32) =      .
                                                        Prob > chi2   =      .
Log pseudolikelihood = -225.15344                       Pseudo R2     = 0.2432
```

startu~O	dF/dx	Robust Std. Err.	z	P>|z|	x-bar	[95% C.I.	]
related*	-.0010603	.0272395	-0.04	0.969	.084433	-.054449	.052328
CVC_li~d*	-.0582116	.0444127	-1.58	0.115	.825858	-.145259	.028836
log_st~g	.0494762	.0117083	4.02	0.000	3.31259	.026528	.072424
log_CV~y	.0112133	.0099916	1.12	0.264	1.21609	-.00837	.030797
log_ro~r	-.0074954	.0225978	-0.33	0.741	1.08397	-.051786	.036795
log_ea~d	.0096312	.0189885	0.51	0.613	.782279	-.027586	.046848
log_to~s	-.0109857	.0216597	-0.51	0.611	2.00798	-.053438	.031467
log_ag~g	.0252546	.0124558	2.07	0.039	3.95153	.000842	.049668
CVC_le~r*	-.0448309	.0298896	-1.02	0.308	.100264	-.103413	.013752
log_CV~s	.0012224	.0100512	0.12	0.903	3.87938	-.018478	.020922
_Ii~2003*	-.0094289	.0597372	-0.15	0.881	.043536	-.126512	.107654
_Ii~2004*	.1316822	.1232765	1.48	0.140	.056728	-.109935	.3733
_Ii~2005*	.1753437	.1481343	1.70	0.090	.042216	-.114994	.465682
_Ii~2006*	.1342852	.1219686	1.51	0.130	.055409	-.104769	.373339
_Ii~2007*	.0213029	.0705078	0.33	0.741	.077836	-.11689	.159496
_Ii~2008*	-.0009564	.0641106	-0.01	0.988	.077836	-.126611	.124698
_Ii~2009*	.0469519	.0848375	0.66	0.509	.05409	-.119326	.21323
_Ii~2010*	.0149486	.0727848	0.22	0.825	.068602	-.127707	.157604
_Ii~2011*	.0488331	.0839291	0.69	0.488	.080475	-.115665	.213331
_Ii~2012*	.0334233	.0770645	0.50	0.619	.081794	-.11762	.184467
_Ii~2013*	.0877293	.0932909	1.21	0.228	.110818	-.095118	.270576
_Ii~2014*	.0362345	.0654119	0.62	0.535	.199208	-.091971	.164439
_I_di~_5*	.9501122	.0149481	5.05	0.000	.038259	.920814	.97941
_I_di~_6*	.9674066	.0182919	5.66	0.000	.097625	.931555	1.00326
_I_di~_7*	.934466	.059312	6.76	0.000	.415567	.818217	1.05072
_I_di~10*	.9674991	.0246547	5.35	0.000	.114776	.919177	1.01582
_I_di~11*	.9439028	.0158462	4.98	0.000	.026385	.912845	.974961
_I_di~16*	.9644191	.0438123	5.57	0.000	.211082	.878548	1.05029
_I_di~17*	.948628	.0127761	5.20	0.000	.031662	.923587	.973669
_I_di~18*	.9393261	.0117	5.25	0.000	.005277	.916395	.962258
_I_di~19*	.9370954	.0182956	4.64	0.000	.015831	.901237	.972954
_I_di~24*	.9597873	.008515	6.71	0.000	.039578	.943098	.976476
_I_di~26*	.9374659	.011975	.	.	.002639	.913995	.960936
obs. P	.1332454						
pred. P	.0643179	(at x-bar)					

(*) dF/dx is for discrete change of dummy variable from 0 to 1
    z and P>|z| correspond to the test of the underlying coefficient being 0

Appendix 5 – FF12 industry classification conversion table

FF12 Conversion Table

FF12 Category	SIC Codes
01- Consumer Nondurables -- Food, Tobacco, Textiles, Apparel, Leather, Toys	0100-0999; 2000-2399; 2700-2749; 2770-2799; 3100-3199; 3940-3989
02- Consumer Durables -- Cars, TVs, Furniture, Household Appliances	2500-2519; 2590-2599; 3630-3659; 3710-3711; 3714-3714; 3716-3716; 3750-3751; 3792-3792; 3900-3939; 3990-3999
03- Manufacturing -- Machinery, Trucks, Planes, Off Furn, Paper, Com Printing	2520-2589; 2600-2699; 2750-2769; 3000-3099; 3200-3569; 3580-3629; 3700-3709; 3712-3713; 3715-3715; 3717-3749; 3752-3791; 3793-3799; 3830-3839; 3860-3899
04- Oil, Gas, and Coal Extraction and Products	1200-1399; 2900-2999
05- Chemicals and Allied Products	2800-2829; 2840-2899
06- Business Equipment -- Computers, Software, and Electronic Equipment	3570-3579; 3660-3692; 3694-3699; 3810-3829; 7370-7379
07- Telephone and Television Transmission	4800-4899
08- Utilities	4900-4949
09- Wholesale, Retail, and Some Services (Laundries, Repair Shops)	5000-5999; 7200-7299; 7600-7699
10- Healthcare, Medical Equipment, and Drugs	2830-2839; 3693-3693; 3840-3859; 8000-8099
11- Finance	6000-6999
12- Other -- Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment	

Source: Kenneth R. French (2020)

Appendix 6 – Robustness test with Fama French 12 based relatedness calculation

```
Probit regression, reporting marginal effects           Number of obs =    823
                                                        Wald chi2(30) = 121.95
                                                        Prob > chi2   = 0.0000
Log pseudolikelihood = -231.60513                       Pseudo R2     = 0.2441
```

startu~O	dF/dx	Robust Std. Err.	z	P>|z|	x-bar	[95%	C.I.]
related*	-.0069761	.0188559	-0.36	0.716	.277035	-.043933	.029981
CVC_li~d*	-.0384836	.0336696	-1.32	0.188	.804374	-.104475	.027507
log_st~g	.0433906	.0092685	4.54	0.000	3.24907	.025225	.061557
log_CV~y	.0081738	.0085898	0.94	0.346	1.17979	-.008662	.02501
log_ro~r	-.0098805	.0206344	-0.47	0.635	1.07414	-.050323	.030562
log_ea~d	.0079054	.0172336	0.46	0.648	.775275	-.025872	.041683
log_to~s	-.0079187	.0182488	-0.44	0.663	1.97775	-.043686	.027848
log_ag~g	.0252344	.0112827	2.28	0.023	3.96984	.003121	.047348
CVC_le~r*	-.0365194	.0298254	-0.90	0.370	.104496	-.094976	.021937
log_CV~s	.000632	.0083552	0.08	0.940	3.83014	-.015744	.017008
_Ii~2003*	-.023219	.0424859	-0.46	0.646	.044957	-.10649	.060052
_Ii~2004*	.0810638	.0926492	1.14	0.253	.053463	-.100525	.262653
_Ii~2005*	.1097417	.1128886	1.33	0.183	.047388	-.111516	.330999
_Ii~2006*	.0821828	.0915792	1.17	0.243	.053463	-.097309	.261675
_Ii~2007*	-.0068252	.0471235	-0.14	0.889	.077764	-.099186	.085535
_Ii~2008*	-.0147797	.0468227	-0.29	0.775	.076549	-.10655	.076991
_Ii~2009*	.0128377	.0581959	0.24	0.814	.054678	-.101224	.1269
_Ii~2010*	-.0061599	.0534125	-0.11	0.912	.066829	-.110846	.098527
_Ii~2011*	.0216204	.0613438	0.39	0.698	.085055	-.098611	.141852
_Ii~2012*	.0020958	.0518079	0.04	0.967	.085055	-.099446	.103637
_Ii~2013*	.0473076	.0678753	0.83	0.407	.111786	-.085726	.180341
_Ii~2014*	.011884	.0491577	0.25	0.800	.18955	-.084463	.108231
_IFF1~_2*	-.0408239	.0273802	-0.96	0.336	.030377	-.094488	.01284
_IFF12~3*	-.025908	.0272713	-0.81	0.418	.153098	-.079359	.027543
_IFF12~5*	.0671175	.0564412	1.50	0.133	.109356	-.043505	.17774
_IFF12~6*	.2481248	.1635587	2.34	0.019	.042527	-.072444	.568694
_IFF12~7*	-.0120953	.0323056	-0.35	0.723	.194411	-.075413	.051222
_IFF12~8*	.0023662	.0570252	0.04	0.966	.029162	-.109401	.114134
_IFF12~9*	.0317436	.0765287	0.49	0.627	.023086	-.11825	.181737
_IFF1~10*	.1303829	.0502761	3.48	0.001	.291616	.031844	.228922
obs. P	.1227217						
pred. P	.0597255	(at x-bar)					

(*) dF/dx is for discrete change of dummy variable from 0 to 1
z and P>|z| correspond to the test of the underlying coefficient being 0