LUISS

Department
of Economics and Finance

Course of Economics and Business

Research & Development: An industry oriented analysis of stock market returns

Domenico Visconte

Prof. Ugo Zannini

SUPERVISOR

IDNo.215861

CANDIDATE

Academic Year 2018/2019

To my family,
to my friends,
to myself.

Table of Contents

- Front page......i
- Acknowledgments......ii
- Table of contents......iii

1. Introduction......1
2. Intangible assets......4
   - 2.1. Overview......4
     - 2.1.1. Accounting treatment and descriptive statistics......12
   - 2.2. Research and Development......15
     - 2.2.1. Accounting treatment and descriptive statistics......16
     - 2.2.2. Literature review......21
3. The empirical model......24
   - 3.1. Data and sample......25
   - 3.2. Model construction......25
     - 3.2.1. Fama and French: The 5-factor model......26
     - 3.2.2. Model and variables construction......30
   - 3.3. Results and implications......37
4. Conclusion......42

- Bibliography......43
- Appendix A......46
- Appendix B......47

Chapter 1

Introduction

For centuries, the economic value of an asset has been estimated through a process of counting and measuring physical goods, especially those with lasting value. Eventually, professionals grouped these assets under the balance sheet's headings of fixed assets or investments. Yet, as most countries developed their respective economies, the nature of investment began to take on a whole different meaning than the one captured and intended by applicable accounting principles. The inadequacy of accounting rules was to some extent already acknowledged during the 60s; nonetheless, as the Internet and computers became ever more rooted in the society, the idea of intangible assets being important drivers of economic value gained wider consensus.

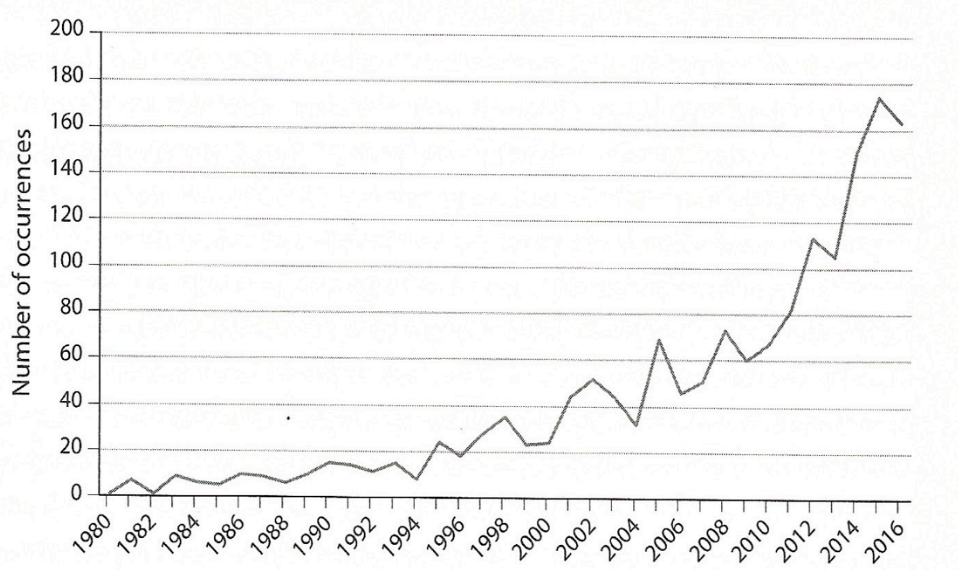

Fig 1: Number of mentions of the word "intangible" in the Title, Abstract or Keyword in academic journals in the field "Economics, Econometrics and Finance" recorded in the database Science Direct.[1] Source: Capitalism without Capital.

From Fig.1 it is possible to appreciate the trend of intangible assets' public awareness. The graph does proxy the latter variable by the number of mentions of the word *intangible* in academic journals. As a matter of fact, professionals began to devote their efforts towards the study of intangible assets, in this way creating the extensive literature that exist in current days. A great contribution to the development of measurement frameworks and real world data was given by

[1] Haskel, J., & Westlake, S. (2018). *Capitalism without capital: the rise of the intangible economy*. Princeton, NJ: Princeton University Press.

Corrado, Hulten and Sichel (2005) in which they published estimates of how much US businesses invested in intangible assets, along with the methodology used to extrapolate those figures. Supported by international organizations and national institutions, several studies produced extensive data on intangible investments that led to the introduction of some types of intangibles, particularly R&D, in investment surveys by statistical agencies.

What may have just seemed as a pundits' debate trend was instead a real phenomenon advancing at unrelenting pace. Sure enough, Fig.2 presents both tangible and intangible investment as a share of the United States and Europe's aggregate Gross Domestic Product (GDP) over time.

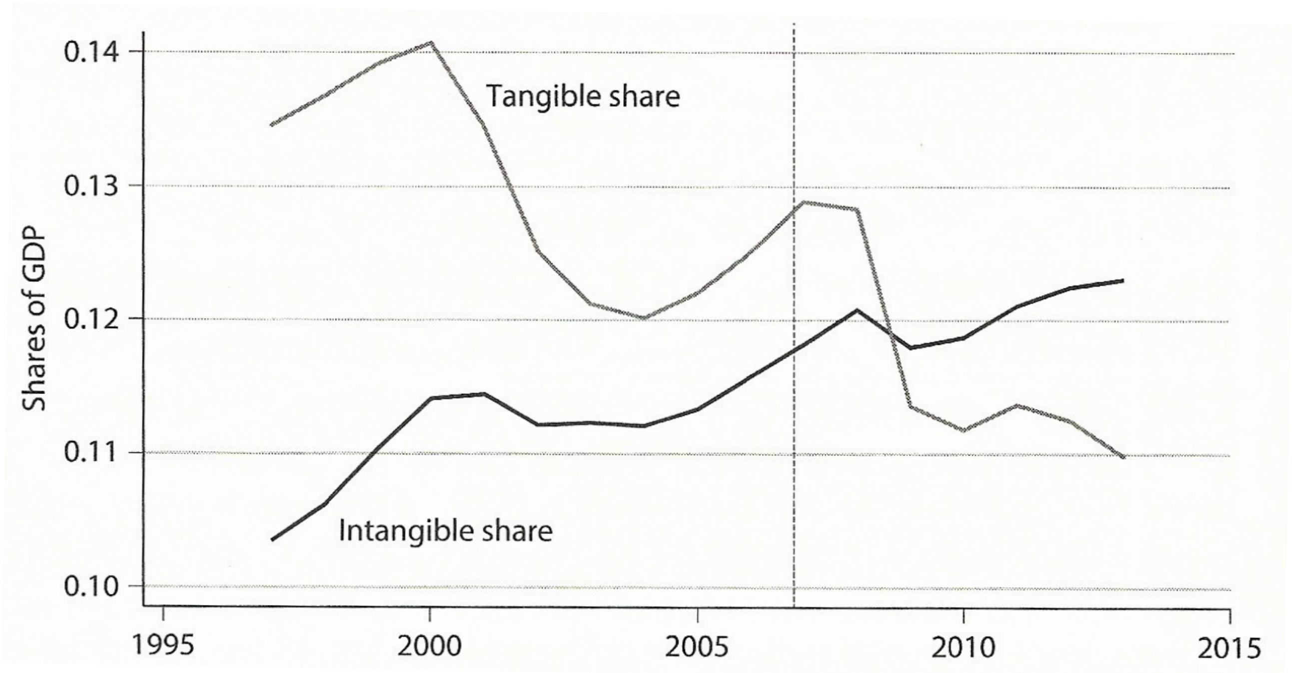

Fig.2: Intangible and Tangible investment in Europe and the United States.[2] *Source: Capitalism without Capital.*

The graph represents a modern reconstruction of past investments, inclusive of unreported intangible assets, showing the evolution of aggregate investment. As this trend evolved, both economies experienced a major structural shift. It is estimated that intangible investment surpassed tangible ones in 2009. Nowadays, even though accounting standards do not still account for most intangibles, they are recognized as a key driver of modern societies and their representation in world economies have never been higher.

Being drivers of value, intangible assets' current relevance in the modern world undeniably inspired the scope and purpose of this dissertation. The study focuses on a specific type of intangible assets: Research and Development (R&D). More precisely the focus of the dissertation is that of exploring

[2] Haskel, J., & Westlake, S. (2018). *Capitalism without capital: the rise of the intangible economy*. Princeton, NJ: Princeton University Press.

the relationship between financial markets valuation of the intangible resource and the industries' commitment to the underlying asset. Because, common sense suggests that the context of application of any asset is a critical determinant of its value we opt for a study that eliminate this dependency. Accordingly, we focus on the market valuation of R&D asset for twenty-four distinct industries so that, in principal, the results would reflect the proposed idea relative to application context independency.

The study follows a precise structure, shaped in order to provide the reader with the knowledge required for understanding, more or less clearly, the implications stemming from the estimated results.

Chapter 2 focuses on both a holistic and a technical definition of the "playing ground". It does so by means of an in depth characterization of intangible assets, in section *2.1*. Next, in subsection *2.1.1* we propose a context dependent definition, one that is shaped by national accounting standards. At the same time the subsection is concluded with a very brief overview of real world figures, with the purpose of providing a snapshot of intangible asset investment's current state. In section *2.2* the discussion narrows the focus to Research and Development. The structure of the latter and *2.2.1* follows the same direction as *2.1* and *2.1.1*, however by placing a greater emphasis on technical details, useful towards the comprehension of the empirical section. Subsection *2.2.2* illustrate the existing literature on the research topic. It is organized in such a way as to communicate how it has evolved, where it currently stands and how critically it has evaluated the produced knowledge.
Chapter 3 explains in detail how the research focus is going to be addressed and evaluates its outcomes. It starts by providing a brief insight on the following sections. Section *3.1* describes technicalities of the data sample. Section *3.2* describes the construction of the empirical model. It splits into two subsections. In *3.2.1* the Fama and French five-factor model is introduced and its construction is explicated. It is covered in great depth as it represents the entirety of control variables employed in the research's final model. In *3.2.2* the final empirical framework is explained. The section contains information on model assumptions, variable and model construction, integration with Fama and French and general model specifications. In section *3.3* results are presented and discussed along with some degree of critical model revision.
Finally, *Chapter 4* covers the conclusion in which a brief summary of findings and implications is laid out, along with recommendations for future research.

Chapter 2

Intangible Assets

2.1 Overview

The word *intangible* is an adjective that refers to the absence of physical matter. When such term is coupled with the economic meaning of the word *asset,* we are essentially linking a physical characteristic to a resource that is expected to provide future benefits to its owner. Therefore, the meaning of intangible asset matches concepts such as software, ideas, knowledge, brands, processes, relationships and many more, to the extent that they provide an economic return, within the context of application.

Moreover, for the sake of clarity we ought to define what is intended by *investment* and, consequently, by *intangible investment*. According to the UN's System of National Accounts an investment *"is what happens when a producer either acquires a fixed asset or spends resources to improve it"*[3]. From the above statements we may logically state that *intangible investment* refers to a producer's expenditure of resources incurred with the intent of creating a long-lived non-tangible asset.

A more in depth definition should include a description of the different economic characteristics stemming from the distinctive physical nature of intangible assets. In fact, the latter exhibit qualities extremely different from those associated with tangible assets. In addition, identifying and describing the unusual qualities is of utmost importance because of the real world implications that these originate, both at a systematic and corporate level.

According to Jonathan Haskel and Stian Westlake's (2018) *Capitalism without Capital*, intangible assets' features can be summarized into four S:

- *Scalability*;
- *Sunkenness*;
- *Spillovers*;
- *Synergies*.

---

[3] Haskel, J., & Westlake, S. (2018). *Capitalism without capital: the rise of the intangible economy*. Princeton, NJ: Princeton University Press.

Before delving deeper into the four S definitions we report an illustrative example taken from the aforementioned book in an attempt to shed some light on their dynamics.
The example recounts a business project of Electric and Musical Industries Limited (EMI), historical owner of Parlophone, record company of the iconic English rock band, The Beatles. EMI was not only involved in the music industry. Among its products it marketed commercial computers, kettles, guided missiles and other electric related goods. Backed by huge cash flows generated from The Beatles' success, the firm was able to invest in the development of the first computed thermography scanner (CT). The technology allowed doctors to make accurate 3D representations of patients' soft tissues. It received wide acclamation to the point that it earned a Nobel Prize and a knighthood to the lead researcher behind the breakthrough. Yet, CT scanners were a commercial failure that ended with EMI licensing the technology to dominant competitors, only to later exit the market entirely.

The illustrative example offers four basic insights:

i. Beatles' music yielded high enough return to back a capital-intensive business such that of CT scanner. Music rights fall in the realm of intangible assets; once generated it can be used over and over again to produce additional units of the final product at virtually zero or minimal costs. Intangibles are scalable.

ii. When EMI decided to set up the CT scanner business, most of the expenditures were directed towards creating the R&D to design the apparatus, training clinicians on how to use the technology and building a nationwide recognised brand by investing in marketing activities. Despite profitably licensing the technology thanks to its scalability, the company opted to exit the market. Most of the above investments are firm specific and as such are irrecoverable. In most occasions intangibles are sunk.

iii. Even though EMI had a first mover advantage other firms such as General Electric and Siemens eventually won dominant position in the market. These two firms couldn't have taken advantage of EMI's facilities to produce CT scanners but instead could appropriate the benefits spilling over the intangible investment made by EMI. Essentially, the latter created a market for GE and Siemens. For this reason intangibles are said to produce spillovers effect.

iv. EMI's central R&D lab was involved in many different businesses within the electrical engineering industry. Combining the laboratory's knowledge together with the clinical expertise of Hospital promoted a medical breakthrough. Indeed, investments in intangibles generate synergies; a combination of non-physical assets is worth more than the sum of the individual ones.

Several researches offer enlightening insight with respect to the four key features of intangible assets. Among the contributions to the literature around intangibles we have Arrow (1962) that provides a mathematical formalization of the dynamics of spillovers effects between firms within the same industry and provides the foundations for the extension of Romer (1990). In turn, built on the two previous literature contributions, Glaeser (2011) addresses and demonstrates the significance of spillovers across industries, and Arthur (2009) works on the impact of synergies by emphasizing the economic relevance of combining differing types of knowledge.

We now investigate intangible assets' distinctive characteristics.

- *Scalability*

The word denotes an intangible's ability to be used over and over again in order to produce an arbitrary large number of units. Once an intangible is created there is no need to replicate it for the production of each single good. For instance, once McDonald drafts an operating manual, the only investment that needs to be made in order be useful on a global scale would be that associated to translating the content into the required languages. When that is accomplished, any McDonalds' operating unit will use the manual. In contrast a furniture business will produce single goods that can only be enjoyed by one user; in this sense the latter are rivalry goods.

In an economy whose assets composition is becoming ever more focused on intangibles we may expect to encounter new market dynamics. A simple posteriori analysis can already introduce us to some of the changes that are currently taking place in world economies.

The high tech market usually stands to comprise several industrial sectors whose firms, as the name suggests, make intensive use of technological assets, mostly recognized as intangible assets under international accounting principles. According to Brand Finance™ Microsoft Corp., Alphabet Inc. and Facebook Inc. own a total intangible value corresponding to 95%, 85% and 89% of their enterprise value, respectively.

Rank 2018	Rank 2017	Rank +/-	Company	Sector	Total Intangible Value (USD bn)	Total Intangible Value/ Enterprise Value (%)	Tangible Net Asset Value (USD bn)	Net Disclosed Intangibles (USD bn)	Disclosed Goodwill (USD bn)	Undisclosed Intangible Value (USD bn)	Enterprise Value (USD bn)
1	3	2	Amazon.com Inc	Internet & Software	827	96%	34	3	13	810	861
2	2	0	Microsoft Corp	Internet & Software	686	95%	89	10	35	641	775
3	1	-2	Apple Inc	Technology & IT	648	85%	264	2	6	640	913
4	4	0	Alphabet Inc	Internet & Software	576	80%	239	3	17	556	815
5	12	7	Alibaba Group Holding	Internet & Software	478	93%	46	2	18	458	524
6	6	0	Facebook Inc	Internet & Software	470	89%	101	2	18	450	572
7	11	4	Tencent Holdings Ltd	Internet & Software	443	91%	44	3	4	437	486
8	9	1	Johnson & Johnson	Pharma	351	101%	-5	53	32	265	346
9	5	-4	AT&T Inc	Telecoms	311	88%	42	114	105	91	353
10	7	-3	Anheuser-Busch InBev	Drinks	308	101%	-2	46	141	121	306

Fig.3: Top 10 companies ranked by Total Intangible Value.[4] Source: GIFT™ 2018.

[4] Brand Finance Institute. (2018). GIFT 2018

Such high levels of intangible capital allow these firms to fully assimilate the feature in order to be themselves, as organization, highly scalable. An economy in which these asset compositions will become increasingly common will also have a growing number of mega corporations. Moreover, the prospects of large markets will push many entrepreneurs to set up a potential new entrant in order to gain a share of the profit pool. At the same time these markets will exhibit serious competition from already grown and scalable businesses. Hence we might expect to see highly concentrated market served by large dominant corporations. Finally, in markets where the final product itself is theoretically infinitely scalable (i.e. Google's search engine), we might also expect a scenario where the one firm that wins a dominant position because of superior product performance enjoys almost the entire market being able to service any number of customers.

- *Sunkenness*

When a firm sets up and commits to an investment plan there will be some cost that are not recoverable if the firm decides to opt out of it. Sunkenness does exactly refer to the idea of being unable to get back all or part of the investment by selling the asset created until the reversal decision. These are called sunk costs.

Sunkenness earns its place among the key attributes of intangible assets because is extremely more entangled with the nature of the latter than it is with tangible ones.
Because the attribute is not exclusive to the world of intangible, in order to illustrate the concept and the degree of an asset's sunkenness we prefer to visualize a line representing the spectrum of all assets. At the left extreme we find resources whose recoverability is highly probable and whose prices are relatively stable among various transactions while at the opposite end we find assets whose costs are extremely difficult or even impossible to recover and whose prices exhibit large variation between trades. Even though both types of asset share the feature, the distribution of tangible ones will be concentrated on the left of the spectrum while intangible will tend to be concentrated on the right end.
This tendency owes to two characteristics of intangible (and tangible) that makes them harder to sell. First, intangible assets are usually firm specific and developed internally as opposed to tangible whose mass production and standardization makes them highly marketable. And that is exactly why intangible are most of the time sunk. Lack of standardization and low production volumes are two characteristics that make the creation of secondary liquidity providers market almost impossible, as of the time this text is being written. The absence of secondary market

therefore makes intangible asset's prices an outcome of private negotiations. These are also spurious and hard to set up since there is no such structure servicing the function of bringing buyers and sellers together.

Second, intangible assets are far more likely to be deeply interrelated to the firms that developed them. Firm specificity increases price variability and presents an impediment to the rise of secondary market. The most likely scenario is one in which the asset is overly firm specific for it to be marketable and to be of any value to a potential buyer. Nevertheless, so long as a potential match between acquirer and seller exists, an asset's worth will vary depending on its contribution to the value of the former business. Thus, depending on the two parties, different negotiations will inevitably result in different valuations.

Some new phenomena will arise as result of an economy in which irrecoverable costs represent an increasing proportion of total costs.

Banks will have major issue when it comes to financing intangible intensive businesses. Usually, a financial institution will require some form of collateral to serve as an insurance against the default of the borrower. In that case, the bank will sell the collateral to recover part of the borrowed funds.

The first problem would be that of failing to separate the intangible from the developer's business. In some cases, for example, the know-how of a business cannot be sold because it is found within employees' heads. In others, legal property rights do to some extent diminish the magnitude of the impediment for some types of innovation. In the former case the bank would have nothing to use as collateral to the loan.

However, even when the first issue is dealt with, a second problem would be that of pricing the intangible asset. Absent a secondary market, information on prices cannot flow properly through the industry and in turn a general consensus of equilibrium values cannot be achieved. In this case, the bank would need to either seek a specialized service or devote some corporate resources toward the negotiations of prices.

Thus, today's banking processes are not able to smoothly handle the financing of small intangible intensive business. If current financing dynamics would be used to serve these businesses, the economy itself would experience higher level of systematic risks, therefore becoming more vulnerable. New ways of funding (i.e. venture capital) may suit the characteristics of companies investing in intangible.

- *Spillovers*

Tangible assets are regulated by well-understood property rights consisting of a solid legal framework being the outcome of more than four thousand years of human experience. While the first rules on the ownership of tangible assets date back to 2000 B.C. Mesopotamia, a first attempt to regulate ownership rights of intangibles was probably made in the late medieval Venetian period when written laws protected glass-making techniques. As it is clear, the debate on property rights of intangible assets is an extremely recent one that will need more time and conflicts of ideas in order to produce regulations functioning as smoothly as those regarding tangible assets.

Thus, the nature of intangible investment, that of being non rival and non excludable, combined with blurred regulations creates an environment in which the benefits originating from such highly scalable investments may produce benefits that are rarely captured in full by the investor. This phenomenon is known as spillovers effect.

The industry wide adoption of iPhone design is one example. Apple's investment in marketing, design and supply chains was enjoyed by the entire industry. Competitors all choose smartphone's design that, within legal boundaries, resembled iPhone's aesthetic in order to benefit from the hype and attractiveness generated by Apple's investment.

Another example is that of McKinsey & Co. innovation in organizational design and training. They developed a new hiring strategy that aimed at bringing into the firms graduates from elite colleges to let them work in teams. This, along with a bullish culture of performance and reward attracted exactly the applicants the firm looked for. This new structure that ensured high working rates within the firm, has become the norm in the modern consulting business environment.

The existence of spillovers has three key implications.

First, in an environment where an investor is not sure to obtain the benefits in spite of good performance will undoubtedly invest less. We could visualize the effect by picturing the distribution of the returns associated with an investment in an intangible asset. With some probability spillovers may doom good outcomes and turn them into bad results, essentially shifting some of the mass above the mean to the opposite side, reducing the attractiveness of the overall investment.

Second, the ability to manage and capture one's spillovers as well as the ability to exploit spillovers from other companies' investment translates into superior performance. In a way, we might describe the gain in performance as those enjoyed by a first mover and the early followers in innovative markets. The former usually has some advantages since, absent competition; the

company can earn high margins from a consistent market share that is captured with relatively little effort. A first mover finds its advantages also in the form of reputation gains, better distribution channels and better access to resources. The early followers, instead, are those that should be able to reap consistent returns while contribution to growing the market into a mature one. In addition, superior performance is also compatible with a natural behaviour of individual companies. Maximization of profits in a tangible context is equivalent to the process of minimizing spillovers and maximizing the value of the intangible investment along with the benefits that flow from it.
Third, spillovers have always played a key role in the geographical formation of economies. The effect will be much greater in an intangible rich economy. According to Glaeser (2011) cities are a structure that allows people to capture the benefits of spillovers. He suggests that people seems to have an increased willingness to pay high rents to live in proximity to those doing the same. This puzzle of urbanization may hint at increased spillover benefits of living in cities.

- *Synergies*

Since the beginning of civilization, combining resources allowed us to develop new technologies aimed at enhancing living standards and at promoting our evolution. Synergies, as the above statement's logic implies is not a feature that is exclusive to intangible assets. A stone and a wooden stick enabled us to defend ourselves as much as tort law does these days; they both acted as deterrents against offenders.
Yet, synergies are identified as one of intangibles' distinctive characteristics because their value creating potential is much greater. Some examples are provided by two market successes being AirBnB and Uber. The former offers lodging and tourism experience, while the latter offers city rides and peer-to-peer ride sharing. These are centuries old businesses that over time earned on average equilibrium returns in their respective markets and yet these two companies now have staggering valuations and turnover billions of dollar each year. What has changed is found in a combination of IT technology, networking devices and the knowledge of the standard business in which the former two are applied. Thus, Uber and AirBnB created massive value by investing in new business models (Uber doesn't own any car and AirBnB doesn't own any real estate) and in building big networks, as well as nurturing them by exploiting the so-called network effect. Thus, intangible can be combined among them and/or with tangible in order to create other assets whose application may even be unrelated, spanning across different domains.

The requirement to achieve the value-enhancing potential of synergies is in clear contrast with the risk posed by spillovers. While synergies require a sharing economy in which organizational boundaries should be flexible enough to provide better flows of knowledge and ideas, spillovers hint to the opposite, that is preventing any information leakage and trying to protect corporate ideas both through legal means and by managing resources strategically.

This represents a dilemma for intangible-rich firm. Closing the organization boundaries can prevent spillovers and loss of value but at the same time forecloses any potential synergies between the firm's stock of intangibles and the considerable number of other businesses ideas.

Finally, the combinatorial nature of synergies implies that, in this context, value creation is promoted by trust, cooperation, interdisciplinarity and, among other things, casual exchanges between people.

Now that we have a clearer definition of intangibles assets we may briefly hint at two more characteristics that result as a by-product of the four S we just discussed. These are:

- *Uncertainty*
- *Contestedness*

The performance of intangible investment is considerably more unpredictable than that of tangibles. As pointed before, the four features are not exclusive to the intangible world; they are just exhibited to a much greater degree.

In fact uncertainty is a characteristic not only shown by assets, but one governing the world. Yet it is a worth candidate when it comes to describing intangible assets because, compared to tangibles, the former exhibit a higher downward risk (*sunkenness*) and a higher upward potential (*scalability and synergies*). This tendency to amplify good and bad performances is complicated by spillovers, which do essentially change the distribution of returns. Besides, synergies and spillovers do also contribute to the likelihood of the assets being the object of dispute, an already present dynamic, consequence of ambiguous property rules.

2.1.1 Accounting treatment and descriptive statistics

In section 2.1 we gave a definition of intangible asset aimed at delivering its essence, structured without regard to national and international accounting principles. However, in the real world intangibles assets have several distinct and yet similar definitions, one for each accounting codification. The purpose of proposing new definitions in this section is that of having a technical characterization of what an intangible is. We consider the apparent redundancy a necessity in order to understand real world figures, based on national and international definitions given by accounting standards, hence to have a better grasp of the inputs at the base of financial research, and quantitative analytical methods. In this section we will be discussing intangible assets through a brief comparison between the US GAAP, United States' accounting principles, and the IFRS 3, internationally adopted accounting standard developed by the International Accounting Standard Board (IASB). Under the IFRS 3 (US GAAP adopts different terminology but same practical results) an intangible asset is "*an identifiable non-monetary asset without physical substance." "To meet the definition of an intangible asset, an item lacks physical substance and is:*

- *Identifiable;*
- *Non-monetary; and*
- *Controlled by the entity and expected to provide future economic benefits to the entity – i.e. meets the definition of an asset.*

An intangible asset is 'identifiable' if it:

- *Is separable – i.e. is capable of being separated or divided from the entity and sold, transferred, licensed, rented or exchanged either individually or together with a related contract, asset or liability; or*
- *Arises from contractual or other legal rights, regardless of whether those rights are transferable or separable from the entity or from other rights and obligations."*[5]

In addition to the criteria that an asset must satisfy in order to be considered one, the IFRS 3 develops additional criteria for an intangible to be financially recognised. Accordingly, "*an intangible asset is recognised when:*

- *It is probable that future economic benefits that are attributable to the asset will flow to the entity; and*
- *The cost of the asset can be measured reliably."*[6]

---

[5] 5 KPMG IFRG Limited. (2017). *Ifrs compared to US Gaap*. London

The identification and measurability criteria adopted by the two accounting principles yields an improved measure of that reported before 2001, when all intangibles were grouped under *Goodwill*. As of now, IFRS 3 categorises intangible assets in five different categories:

- Marketing related
- Customer related
- Contract based
- Technology based
- Artistic related

The intangibles are therefore classified accordingly and, as a general rule, initially recorded in the balance sheet at cost. US GAAP, instead, does not establish the same general rule but in practice there is little to no discrepancies of initial reported values.

Categories of intangible asset under IFRS 3

Marketing-Related Intangible Assets	Customer-Related Intangible Assets	Contract-Based Intangible Assets	Technology-Based Intangible Assets	Artistic-Related Intangible Assets
Trademarks, tradenames	Customer lists	Licensing, royalty, standstill agreements	Patented technology	Plays, operas and ballets
Service marks, collective marks, certification marks	Order or production backlog	Advertising, construction, management, service or supply contracts	Computer software and mask works	Books, magazines, newspapers and other literary works
Trade dress (unique colour, shape, or package design)	Customer contracts & related customer relationships	Lease agreements	Unpatented technology	Musical works such as compositions, song lyrics and advertising jingles
Newspapers	Non-contractual customer relationships	Construction permits	Databases	Pictures and photographs
Internet Domain Names		Permits	Trade secrets, such as secret formulas, processes, recipes	Video and audio-visual material, including films, music,
Mastheads		Franchise agreements		videos etc.
Non-competition agreements		Operating and broadcast rights		
		Use rights such as drilling, water, air, mineral, timber cutting & route authorities		
		Servicing contracts such as mortgage servicing contracts		
		Employment contracts		

Fig.4: Classification of intangible assets according to IFRS 3. Source: GIFT™ 2018.[7]

Even though accounting standards enhanced the quality of disclosed financial statements through classification and recognition of a broader set of asset, undisclosed intangible still account for a consistent share of a firm's enterprise value. As a matter of fact, the aggregate measures provided by Brand Finance estimate that, as of the beginning of 2018, global enterprise value stood at $109.3 trillions, of which $57.3 trillions (52%) were accountable to total intangible assets which in turn comprised $43.7 trillions of undisclosed intangible value, representing 40% of global enterprise value. Undisclosed intangible assets are a result in part owing to the employed identification criteria and in part due to the lack of recognition of internally generated asset.

[7] Brand Finance Institute. (2018). GIFT 2018

A further distinction between acquired and internally generated intangible should be made. Under US GAAP and IFRS 3 acquired intangible assets are recorded at fair value. These can be part of a business combination or of a separate acquisition. In the first case, the IFRS 3 assumes that the criteria are met while in a separate acquisition it requires for the analysis under the probability and reliability of measurement criterion. Differently, the US GAAP recognises an intangible asset, being part of a business combination, only if the identification requirements are satisfied while it always recognises an intangible acquired outside of a business combination on the basis that the negotiation itself is proof of the existence of the intangible.
With regards to internally generated assets, the former recognises them only if a specific "Codification subtopic" requires their recognition. In this case their initial value is the cumulative costs incurred after the capitalization criteria (differing for each subtopic) are satisfied. Instead, IFRS 3 records the cumulative costs incurred for the preparation of the underlying asset for its intended use, following the same principles applying to Property, Plant and Equipment.

Despite the improved informative content of financial statements, undisclosed intangibles, in the last five years, still accounts for an average 37,4% of enterprise value while the disclosed value of disclosed Goodwill and intangible assets hover at around 8-10% and 6-8%, respectively.

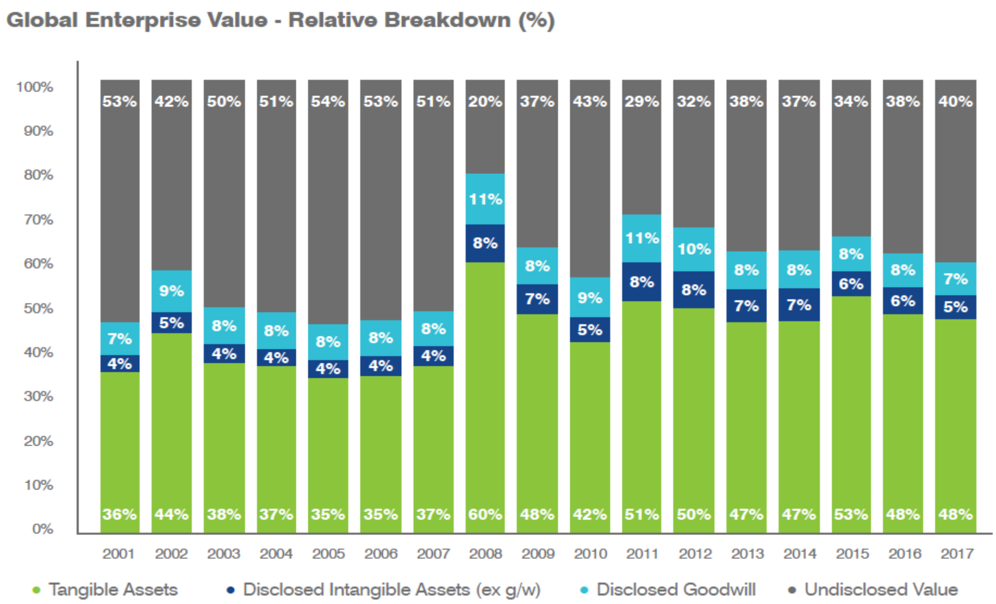

Fig.5: Global composition of enterprise value.[8] *Source: GIFT*^{TM} *2018.*

According to Brand Finance *GIFT*^{TM} *2018,* the scenario has worsened over the past years. They estimate that enterprise value grew by 18% over the 2016-2017 period. At the same time undisclosed value increased by 25% while disclosed goodwill and disclosed intangible assets experienced an aggregate change of 16%.

[8] Brand Finance Institute. (2018). GIFT 2018

2.2 Research and development

The potential contribution of R&D to economic growth and prosperity has been long acknowledged. Over the years, numerous resources and efforts by researchers and, national and international organizations have been devoted to the study and advancement of R&D knowledge. Among the most important contributions to the R&D field of study, the *Frascati Manual* is recognised as a world standard, providing *"the basis for a common language for talking about R&D and its outcomes"*[9]. First drafted in June 1963 by experts from the Organization for Economic Cooperation and Development (OECD) and the National Experts on Science and Technology Indicators (NESTI), it offers a framework for *"collecting and reporting internationally comparable statistics on the financial and human resources devoted to research and experimental development"*[10].

According to the Frascati Manual R&D is defined as an activity comprising a *"creative and systematic undertaken in order to increase the stock of knowledge [...] and to devise new applications of available knowledge"*[11]. In order for an activity to be classified as Research and Experimental Development (R&D) it must exhibit the following characteristics:

- novel
- creative
- uncertain
- systematic
- transferable and/or reproducible

From the criteria, it follows that to be classified as R&D, an activity has to be directed at new findings, which must not be obvious but instead based on original concepts to improve current knowledge. The costs, of time and capital, as well as the final outcomes associated with R&D cannot be accurately predicted; therefore, the activity has to be carried out in an organized manner in order to keep ordered records on both the applied methods and the outcomes achieved. Lastly, the codification of knowledge is crucial for an activity to be classified as R&D. It serves as a way to share the newly created knowledge and promotes its reproduction and in this way really contributed to the improvement of existing *stock of knowledge*.

9, 10, 11 OECD (2015), *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris, https://doi.org/10.1787/9789264239012-en.

2.2.1 Accounting treatment and descriptive statistics

Both accounting standards we have been discussing so far adopt a clear definition of what R&D is and, in particular they differentiate between the broad, general phases of such activity, namely Research and Development.

According to the IFRS, *Research* is defined as an *"original and planned investigation undertaken with the prospect of gaining new scientific or technical knowledge and understanding"* while *Development* as the *"application of research findings or other knowledge to a plan or design for the production of new or substantially improved materials, devices, products, processes, systems or services before the start of commercial production or use"*[12]. Under US GAAP the former is intended as *"a planned search or critical investigation aimed at the discovery of new knowledge with the hope that such knowledge will be useful in developing a new product or service or a new processes or technique or in bringing about a significant improvement to an existing product, service, process or technique"* and *Development* as the *"translation of research findings or other knowledge into a plan or design for a new product, service, process or technique, whether intended for sale or for use."*[13]

Although both accounting standards employ similar wordings, the treatment of R&D expenses shows some divergences. Both require research costs to be expensed when incurred. IFRS capitalizes expenses on internally generated assets, arising in the development phase, from the date on which the company is able to demonstrate the activity's fit to several criteria. On the other hand, US GAAP does offer an exception to capitalization of expenses incurred in the development phase only to specific internally developed computer software and direct-response advertising.
IFRS initially recognises in-process R&D obtained in a business combination at fair value and in-process R&D obtained in a separate acquisition at cost. Following initial recognition the assets is subjected to general principles governing Research and Development. Likewise, US GAAP employs the same principle for initial recognition of acquired R&D but it diverges from IFRS' treatment of post recognition expenses. Regardless of acquisition methodology, in-process R&D is classified as an indefinite lived assets until abandonment or completion of the project and, unlike IFRS, expenses all costs related to it in case of non-abandonment.

The above discussion leads to a well-defined set of R&D measurements. In order to have a better picture regarding the distribution of R&D expenses, we are now presenting several descriptive statistics considering both time and industries as dimensions of the analysis. The following are

[12, 13] KPMG IFRG Limited. (2017). *Ifrs compared to Us Gaap*. London

constructed from a sample of firms corresponding to the entirety of Russell 3000. R&D figures associated with each individual stock are reported according to US GAAP while industries are classified according to the Global Industry Classification Standard (GICS). We will postpone any discussion about technical choice and implications in the next chapter. In addition, the following statistics are performed on a temporal window consistent with the timing of accounting standards reforms.

The raising importance of R&D intangible assets in the business environment can be appreciated by looking at its growth rate. The graph presented below plots the aggregate rate of growth of R&D expenses of all the firms included in the index.

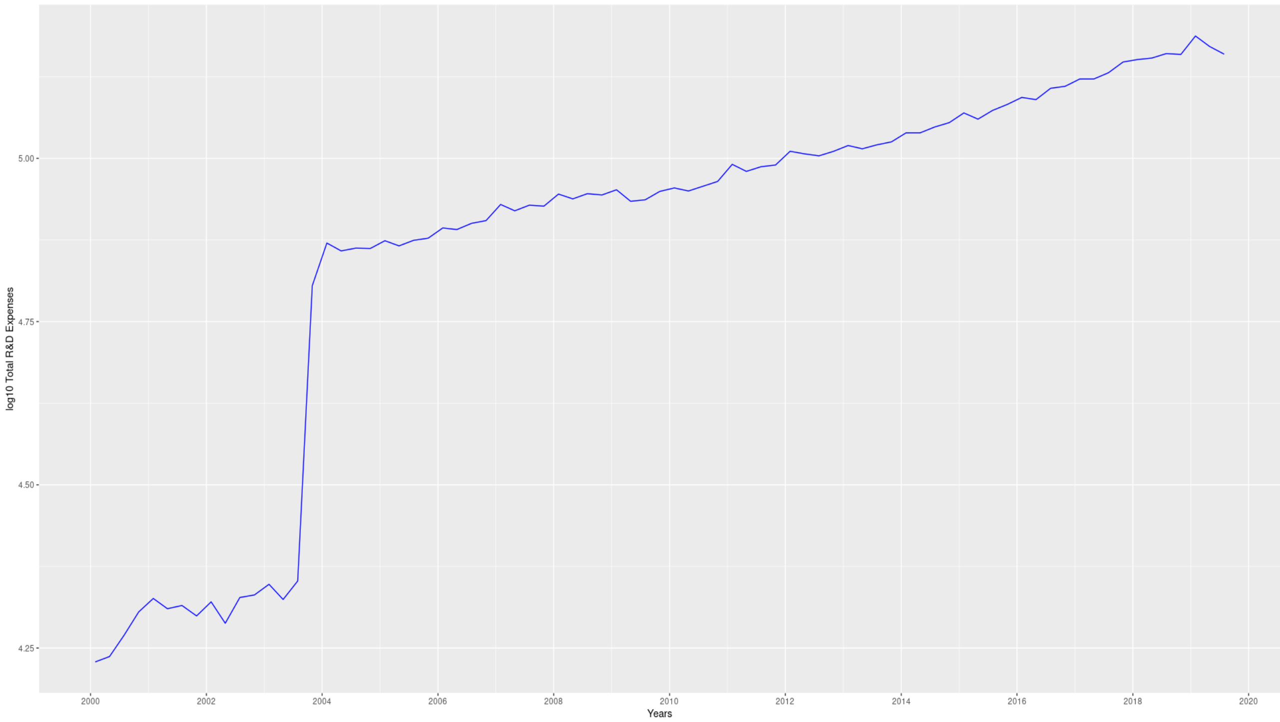

Fig.6: Growth rate of R&D assets over the 2000-2019 period. Source: Author.

From the graph in Fig.6 we notice that over the selected temporal window, the aggregate amount of R&D expenses has steadily increased. In 2003, the massive jump in growth rate is caused both by the increasing number of firms reporting and performing R&D activity, but also because of missing values found in the data that would have otherwise mitigated the effect. Nevertheless, the above is a good representation of how the presence of R&D has evolved over time.

Further examination of the collected sample leads to Exhibit 1.

Exhibit 1. Distribution of R&D Expenditure for Russell 3000 Firms Organized by Four-Digit GICS Industry Group Codes, Firm Averages for 30th April 2001- 31th July 2019 (Dollar Amounts in Millions)

Four-Digit Industry Group	*Industry Group*	*Number of Firms**	*Total R&D ($)*	*Total Asset ($)*	*R&D per Firm ($)*	*R&D/Total Asset*
1010	Energy	107	1024.78	1242548.42	10.15	0.10%
1510	Materials	93	823.57	456292.17	8.49	0.19%
2010	Capital goods	173	37647.37	791478.75	212.34	4.76%
2020	Commercial and professional services	68	137.78	137660.83	1.98	0.10%
2030	Transportation	46	85.59	373862.96	1.59	0.01%
2510	Automobiles and components	14	6118.02	286265.06	414.27	2.21%
2520	Consumer Durables and Apparel	67	303.17	161206.15	4.34	0.18%
2530	Consumer services	69	50.63	280951.31	0.58	0.01%
2550	Retailing	87	2046.05	419436.78	19.59	0.35%
3010	Food & staples retailing	16	0.00	275854.16	0.00	0.00%
3020	Food, Beverage and Tobacco	47	399.14	303682.50	8.06	0.12%
3030	Household & Personal products	17	346.22	56994.94	19.58	0.61%
3510	Health care equipment and services	124	2433.37	636984.46	18.88	0.40%
3520	Pharmaceuticals, Biotech. and Life Sciences**	157	12503.71	625266.11	78.26	2.12%
4010	Banks	2	2.72	1804.88	1.52	0.21%
4020	Diversified financials	20	23.37	476655.64	1.14	0.01%
4030	Insurance	8	0.78	75988.86	0.09	0.00%
4510	Software and services	112	7145.74	566491.84	63.05	1.35%
4520	Technology Hardware and equipment	85	5790.62	518643.97	66.64	1.22%
4530	Semiconductors and Semiconductors equip.**	55	6023.69	236338.82	104.46	2.57%
5010	Telecommunication services	20	416.00	590460.59	18.94	0.06%
5020	Media & Entertainment	62	4318.28	659022.37	54.39	0.47%
5510	Utilities	67	15.56	1030502.03	0.23	0.00%
6010	Real Estate	22	11.65	72653.36	0.61	0.02%

**Number of Firms* is rounded to the nearest integer; **Names are truncated for layout purposes. *Source: Author*

From Exhibit 1 we observe that less than half of the firms in Russell 3000 perform R&D activities. The table reports the number of firms that perform R&D in each industry along with the absolute value of associated total R&D expenses. In addition it provides two measure of intensity being R&D expenses per number of firms and the ratio of R&D expenses to Total Asset. From the latter metric we notice that the reported industries are split into halves. Some industries exhibit ratios greater than 1% while others show mostly null or smaller than 0.4% ratios. High ratio values are found in Capital Goods, Pharmaceuticals, Biotechnology and Life Sciences, Semiconductors and Semiconductors equipment and Technology Hardware and equipment. At the opposite end of the spectrum we find Transportation, Consumer services, Food & Staples retailing, Diversified Financials, Telecommunication services, Real Estate and Utilities. According to these values, we may expect to find positive R&D factors loadings in industries that present a rationale for undertaking the activity. Hence, for R&D efforts to be positively valued only in those industries where the activity is considered to be fruitful and vice versa.

To visually characterize levels of R&D expenses according to industry we employ a logarithmic scale, as in Fig.7. This is employed because, as seen from Exhibit 1, few industries exhibit high levels of R&D expenses while the remaining ones, according to the sample being used, have moderate to inexistent intensities.

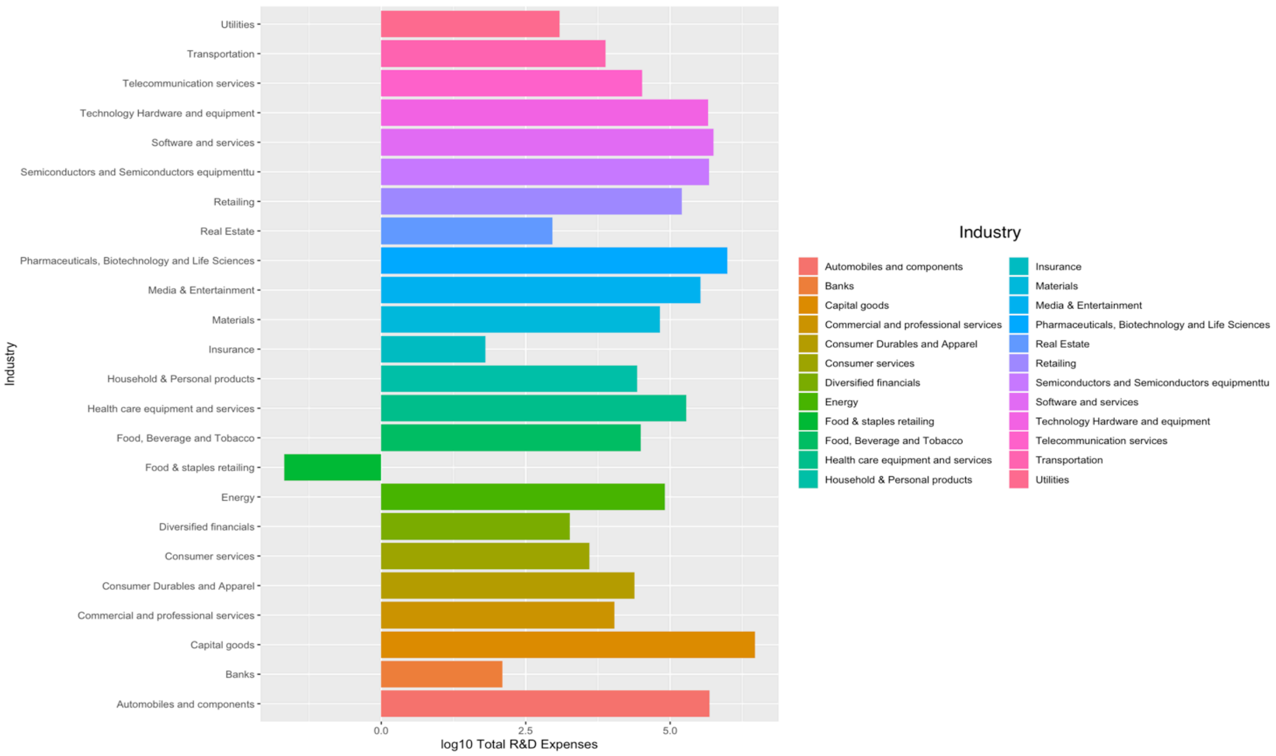

Fig.7: Aggregate values of R&D per industry; logarithmic form. Source: Author.

To better visualize the evolution of R&D and how it became ever more relevant in distinct industries we propose in Fig.8 a graph mapping the growth of the intangible assets over time.

Similarly to previous charts we employ logarithmic scaling in order to identify trends even in those industries that exhibit very low level of R&D expenses.

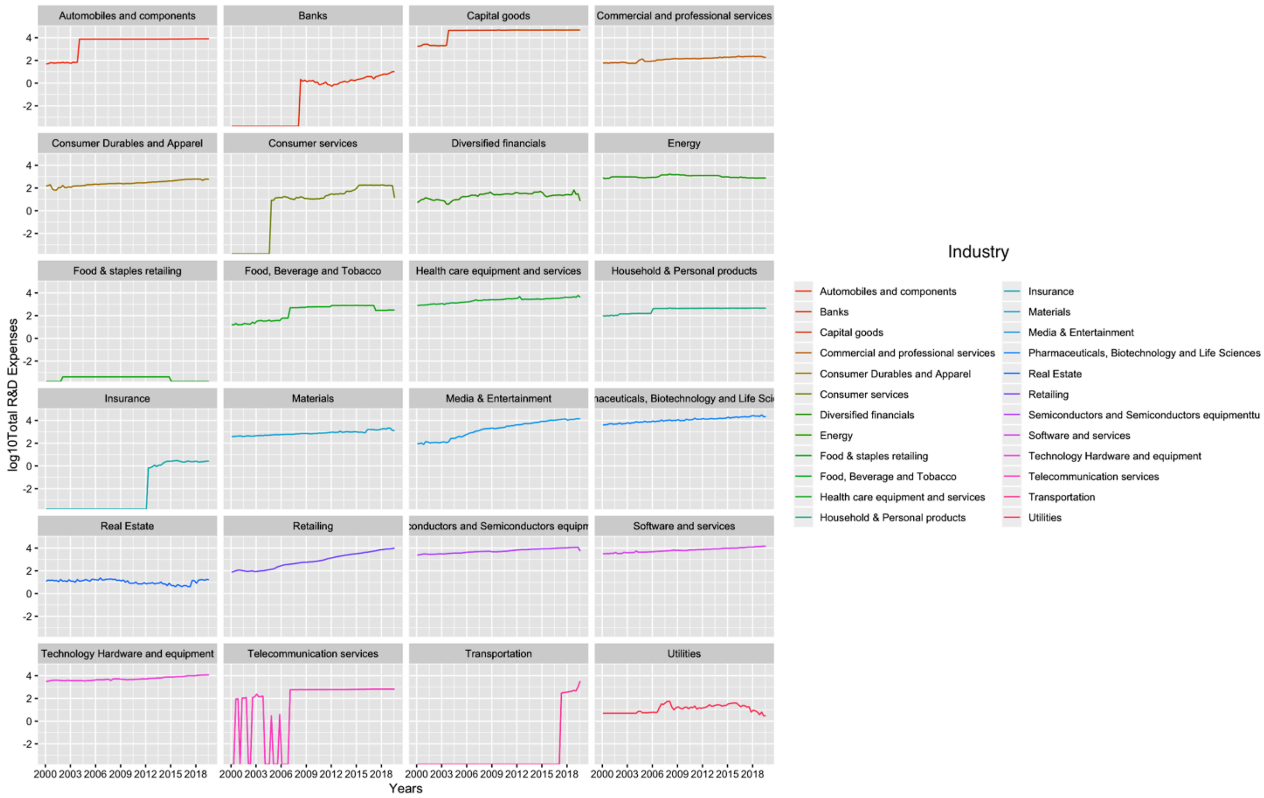

Fig.8: Growth rate of R&D expenses per industry over the 2000-2019 period. Source: Author.

Over the selected temporal period only Utilities exhibit a decreasing rate of growth. Some other industries float around an industry specific value meaning that they alternate period of overinvestment with period of underinvestment in R&D with respect to the average they seem to follow. These are Real Estate, Utilities, Food & Staples retailing, Energy and Diversified financials.

Yet, the graph also shows both upward and downward spikes. This suggests that either the sample collected presents huge holes for some industries making inferences of trends meaningless, or that unlikely, these industry experienced such sharp rise/decline in R&D spending. Nevertheless, instances such as the trend shown for the Telecommunication services industry is clearly due to missing values in the sample.

2.2.2 Literature review

The effect of R&D spending on stock price movements has been extensively documented over the years. Bublitz and Ettredge (1989), Lev and Sougiannis (1996), Edvinsson and Malone (1997), Lev and Sougiannis (1999), Chan, Lakonishok and Sougiannis (1999) found a positive and significant effect between market value of companies, or equivalently stock returns, and R&D expenditure. In particular, Shuangling, Guohua and Lijuan (2018*)* propose a dynamic continuous model to analyse the long-documented relationship between firm value and R&D investment. The model offers insights on the relationship when incorporating tangible assets, risk management and financing. They find that firm value is sensitive to different rates of R&D's obsolescence and that intangible investment, specifically R&D, is much more susceptible to financing frictions than tangible investments do.

A broad literature of event studies confirms the related findings obtained through the quantitative and direct estimation of the factor loadings associated with R&D expenses and related variables. This methodology aims to assess the presence of the R&D-firm value relationship by investigating the returns generated by particular stocks around a specific event date. Jarrell, Lehn and Marr (1985) and Woolridge (1988) document a positive reaction of markets after announcements of increased R&D efforts leading to abnormal returns. This effect is further investigated in subsequent years when Chan, Martin and Kensinger (1990), Doukas and Switzer (1992), Chauvin and Hirschey (1993) find that the magnitude of the effect of announcement of increased R&D spending on stock return increases with firm size.

Several authors also provide evidences of potential flows in the literature concerned with evaluating the extent to which financial markets are able to discriminate innovativeness as a source of value. These are focused with the reliability/unreliability of market agents' expectations to evaluate a real correlation between R&D efforts and profitability. Hirshleifer, Lim and Toeh (2009) suggest that many investors are unable to clearly account for the effect of past R&D expenditure when evaluating market prices. Statman and Sepe (1989) document a positive market reaction to the abandonment of projects that are expected to perform poorly. Ballester, Garcia-Ayuso and Livnat (2003) find that on average investors believe that 88.2% of current R&D expenditure is going to provide future benefits beyond the year of recognition.

Some other researches do not question the ability of financial markets but instead inspect the suitability of R&D expenditure alone as a predictive factor of firm performance and stock price movements. Baker and Freeland (1975) argue that R&D expenditure does not differentiate between failed attempts and successful projects, intended as the development of a new technology or products. Hirschey, Richardson and Scholz (1998) claim that the late sixties and early seventies showed high valuation of R&D activity by financial markets even though it did not exhibit much constancy over time. Hall (1993a, 1993b) provides evidence that during the eighties the value previously attributed to R&D expenditure had sharply declined.

While existing literature over the years was building the stock of knowledge documenting such relationship, several authors began investigating potential tweaks that could have produced more convincing and enlightening results. These researchers all augmented previous models to incorporate patenting activity and other measures of innovativeness. Hsu and Ziedonis (2013), interpreting patents as an intermediate outcome of project's success, incorporate the number of patents in explaining stock price movements, finding that patenting activity is positively priced especially as a signalling device for relatively young firms. Scherer (1997) argues that there exist few patents that are very valuable while a major fraction of those hold little to no value, implying that the distribution of patent's value is particularly skewed. Trajtenberg (1990) and Harnoff, Narin, Scherer and Vopel (1997) suggest that a better measure could be found in the number of citations received in order to proxy for patent's value. Hall (1998) finds that patents are more informative that R&D expenses and that citation weighted patents are even more so. Yu and Hong (2016) implement a model constructed as a blend of Fama and French 3-factor model to control for factors affecting the overall universe of public corporations and the model of Chen and Zhang (2007) to control for corporate-specific factors. They then augment it by introducing two patent related factors. These are constructed by classifying patents into those representing either exploitation or exploration activity. The former allows a firm to enhance short-term performance through increased efficiency, while the latter helps improving long term performance through the development of new capabilities. They report a positive coefficient on exploitation activity and a negative one on exploration, hinting at financial market's short termism.

Another recent paper that employs new techniques to investigate the association of R&D innovation and stock prices is provided by Coad and Rao (2006). They investigate post innovation performance by considering Tobin's q. They argue that the methodology based on OLS regressions is not suited is not suited for the highly skewed distribution of Tobin's q, and accordingly, they employ a quantile regression technique. In addition, they do not employ R&D and patenting activity measures to investigate the link between innovative activity and firm performance but, using Principal Component Analysis, they create an overall variable of innovativeness aimed at discarding the irrelevant variance of individual measures. Consistent with previous literature they show that financial markets recognize innovative activity, and that its effect on market value greatly varies across the distribution of market values themselves. They claim that the stock market recognises a minimal value in the innovative activities undertaken by low values of Tobin's q, while it places great values for those with high Tobin's q.

The literature built around the stock market performance of firms undertaking Research and Development is wide and spans over several years providing a good picture of how it has evolved. We notice how several researchers devoted resources towards proving the existence of such relationship while, at the same time, being able to question the validity of their findings by assessing the reliability of both involved parties; financial markets, as input providers, and professionals, as the designer of analytical models.

The aim pursued in this paper is that of evaluating the effect of R&D expenses and related metrics on industry-wide stock price performance, after controlling for systematic and firm-specific factors.

Chapter 3

The empirical model

The aim pursued in this paper is that of evaluating the effect of R&D expenses and related metrics on stock price performance, after controlling for systematic and firm-specific factors. In order to characterize this relationship, in this paper we will employ a simplified version of the Fama and French 5-factor model augmented with three R&D metrics. These deviation from the original Fama and French model are applied in order to present a general model which adapts to the available database and whose construction's efforts both allow me to comply with time availability and are commensurate to my personal knowledge. In addition, differently from some previous studies, we choose a diverging approach because, in my opinion, the retrieval of historical values of Fama and French (2014) factors would have not been a suitable strategy when applied to a different observation sample. The study will yield a general model comprising three different specifications, each employing different pairs of metrics.

3.1 Data and sample

The data used comes from two sources: the Bloomberg Terminal and the Federal Reserve Economic Data (FRED) St. Louis website. Historical values required for the construction of the Fama and French model are retrieved from Bloomberg and reported under US GAAP accounting. R&D expenses figures were also obtained through the Terminal. The 3-month Treasury bill rate was retrieved from FRED databases. The analysis is performed on quarterly data from 31 April 2001 to 31 July 2019.

We initially start with the entirety of Russell 3000 by collecting quarterly historical prices of the index. Moving to the index composition it should be noted that, in the available data, the number of firms never actually reaches 3000. For each company in the index we collect quarterly data points of R&D expenditure, Net interest expenses, Market capitalizations, Operating income, Total liabilities and Total assets. R&D expenses data is used to build three related metrics. The first is the proportion of R&D expenses with respect to Total assets. The second is a measure of R&D growth and the third is a measure of acceleration of R&D growth. Because this metrics will be employed in the construction of the general model their presence influences the frequency of observation of each of the three specifications. For this reason *Model 1* and *Model 2* are built on a 72-observation sample while *Model 3* on a 74-observation one since the first R&D metric is null for 31/1/2008 and 31/07/2018. We exclude from the sample those firms whose historical data have more than a year of missing value. The rationale behind the non-exclusion of firms that have less than a year of missing data is that of preserving quantity, assuming that quality will be unchanged when looking at industry aggregate factors. Finally, all data points on returns are in decimal form and historical values employed in the factors' construction are quoted in millions.

3.2. Model Construction

As previously discussed we employ a simplified version of the model of Fama and French (2014). The 5-factor model is able to control for both non-systematic and systematic factors. In this way we are able to better quantify the effect related to Research and Development activities on market valuations. To augment the control model we employ an R&D stock formation relationship taken from Chan, Lakonishok and Sougiannis (1999).

3.2.1. Fama and French: The 5-factor model

Eugene Fama and Kenneth R. French are two of the most acknowledged researchers in the Finance literature. Specifically, in 1993 they constructed a worldwide renowned model also known as 3-factor model.

$$R_{it} - R_{Ft} = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + e_{it} \quad (1)$$

It was designed to capture the effect of firm size and that of Book to Market price ratio after controlling for market movements according to the CAPM model.

In this paper we employ one of their most recent works. The underlying model is a revision of the original 3-factor framework. The authors expanded the model with two additional factors after an extensive revision from peer researchers. In particular, Novy-Marx (2013) identifies a measure that proxies expected profitability, being strongly related to average returns; and Aharoni, Grundy and Zeng (2013) propose a feebler but statistically reliable relation between investment and average returns. Thus, the authors provide a financial rationale behind the introduction of additional factors to the Index model (statistical application of the CAPM). They explain the relationship between the selected variables and average returns through the dividend discount model. It states that the "*market value of a share of stock is the discounted value of expected dividends per share*"[14]:

$$m_t = \sum_{\tau=1}^{\infty} E(d_{t+\tau})/(1+r)^{\tau} \quad (2)$$

Where:

- m_t is the price of a share of stock at time t;
- $E(d_{t+\tau})$ is the expected dividend per share for the period $t+\tau$; and
- r is the Internal Rate of Return (IRR) on expected dividends.

Following Modigliani, Miller (1961) the authors restate (2) to show the relationship between expected return and expected investment, expected profitability and Book to Market ratio. Therefore:

$$M_t = \sum_{\tau=1}^{\infty} E(Y_{t+\tau} - dB_{t+\tau})/(1+r)^{\tau}$$

Where:

- M_t is the price of a share of stock at time t;

[14] Fama, E. F., & French, K. R. (2014). *A five-factor asset pricing model*. Elsevier. doi: http://dx.doi.org/10.1016/j.jfineco.2014.10.010

- $Y_{t+\tau}$ is the earnings at time $t + \tau$ available to total equity; and
- $dB_{t+\tau}$ is the change in total book equity, or equivalently $B_{t+\tau} - B_{t+\tau-1}$, change in investment.

Dividing by Book equity at time t:

$$\frac{M_t}{B_t} = \sum_{\tau=1}^{\infty} \frac{E(Y_{t+\tau} - dB_{t+\tau})/(1+r)^{\tau}}{B_t} \tag{3}$$

Finally, from (3) the authors draw three conclusions:

- Fixing everything except M_t and r, a lower value of M_t, or equivalently, a higher value B_t/M_t implies a higher expected return r;
- Fixing everything except $Y_{t+\tau}$ and r, then a higher value of $E\ (Y_{t+\tau})$ implies a higher value of r;
- Fixing everything except $dB_{t+\tau}$ and r, then a higher growth in $B_{t+\tau} - B_{t+\tau-1}$ implies a lower r.

Thus, following an extensive literature focused on the revision of the 3-factor model (1), Eugene Fama and Kenneth R. French augment the original model with profitability and investment factors. The modification yields the following model:

$$R_{it} - R_{Ft} = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it}$$

Where:
- $R_{it} - R_{Ft}$ is the difference between the return on portfolio/security i at time t and the return on the risk-free asset at time t;
- α_i is the intercept of the regression, theoretically capturing the average return above and below that estimated by regressors;
- $(R_{Mt} - R_{Ft})$ is the difference between the return on a market index at time t and the return on the risk-free asset at time t, also called market premium;
- b_i is the factor loading on the market premium, capturing the sensitiveness of the regressand on the respective factor;
- s_i, h_i, r_i, c_i are the factor exposures of their respective regressors (their purpose is equivalent to that of b_i);
- SMB_t [Small minus Big] is a factor defined as the difference between the average returns earned by small capitalization portfolios and the average returns earned by big capitalization portfolios at time t;

- HML_t [High minus Low] is a factor defined as the difference between the average returns earned by high Book to market ratio portfolios and the average returns earned by low Book to Market portfolios at time t;
- RMW_t [Robust minus Weak] is a factor defined as the difference between the average returns earned by robust profitability portfolios and the average returns earned by weak profitability portfolios at time t;
- CMA_t [Conservative minus Aggressive] is a factor defined as the difference between the average returns earned by conservative investment portfolios and the average returns earned by aggressive investment portfolios at time t; and
- e_{it} is the error term of the regression for portfolio/security i at time t;

Fama and French estimate the 5-factor model on a sample made of all stocks in NYSE, AMEX and NASDAQ available on both CRSP and Compustat databases and having share codes 10 or 11 for the period July 1963-Decemeber 2013. In addition, they sort the sample of firms in distinct portfolio using the NYSE median and breakpoints.

Sort	Breakpoints	Factors and their components
2 × 3 sorts on	*Size*: NYSE median	$SMB_{B/M}=(SH + SN + SL)/3 - (BH + BN + BL)/3$
Size and *B/M*, or		$SMB_{OP}=(SR + SN + SW)/3 - (BR + BN + BW)/3$
Size and *OP*, or		$SMB_{Inv}=(SC + SN + SA)/3 - (BC + BN + BA)/3$
Size and *Inv*		$SMB=(SMB_{B/M} + SMB_{OP} + SMB_{Inv})/3$
	B/M: 30th and 70th NYSE percentiles	$HML=(SH + BH)/2 - (SL + BL)/2=[(SH - SL) + (BH - BL)]/2$
	OP: 30th and 70th NYSE percentiles	$RMW=(SR + BR)/2 - (SW + BW)/2=[(SR - SW) + (BR - BW)]/2$
	Inv: 30th and 70th NYSE percentiles	$CMA=(SC + BC)/2 - (SA + BA)/2=[(SC - SA) + (BC - BA)]/2$
2 × 2 sorts on	*Size*: NYSE median	$SMB=(SH + SL + SR + SW + SC + SA)/6 - (BH + BL + BR + BW + BC + BA)/6$
Size and *B/M*, or	*B/M*: NYSE median	$HML=(SH + BH)/2 - (SL + BL)/2=[(SH - SL) + (BH - BL)]/2$
Size and *OP*, or	*OP*: NYSE median	$RMW=(SR + BR)/2 - (SW + BW)/2=[(SR - SW) + (BR - BW)]/2$
Size and *Inv*	*Inv*: NYSE median	$CMA=(SC + BC)/2 - (SA + BA)/2=[(SC - SA) + (BC - BA)]/2$
2 × 2 × 2 × 2 sorts on	*Size*: NYSE median	$SMB=(SHRC + SHRA + SHWC + SHWA + SLRC + SLRA + SLWC + SLWA)/8$
Size, *B/M*, *OP*, and *Inv*		$- (BHRC + BHRA + BHWC + BHWA + BLRC + BLRA + BLWC + BLWA)/8$
	B/M: NYSE median	$HML=(SHRC + SHRA + SHWC + SHWA + BHRC + BHRA + BHWC + BHWA)/8$
		$- (SLRC + SLRA + SLWC + SLWA + BLRC + BLRA + BLWC + BLWA)/8$
	OP: NYSE median	$RMW=(SHRC + SHRA + SLRC + SLRA + BHRC + BHRA + BLRC + BLRA)/8$
		$- (SHWC + SHWA + SLWC + SLWA + BHWC + BHWA + BLWC + BLWA)/8$
	Inv: NYSE median	$CMA=(SHRC + SHWC + SLRC + SLWC + BHRC + BHWC + BLRC + BLWC)/8$
		$- (SHRA + SHWA + SLRA + SLWA + BHRA + BHWA + BLRA + BLWA)/8$

Fig. 9: Portfolios and factors construction. Source: Fama and French (2014).

They apply three different sorting strategies according to a varied breakpoints scheme for each sort. Accordingly, we now disentangle the first sorting scheme namely, 2x3, in order to clarify the process whose underlying logic is analogous to that applied to the 2x2 and 2x2x2x2 sorts.

The 2x3 sorting strategy entails a two-phase process. The first process involves a sort on firm's size according to NYSE median, basically dividing the entirety of the investable universe into two distinct sections. The second phase entails an additional sorting on B/M, Operating profitability and Investment applied on each of the two previously derived investable universe using as breakpoints the 30th and 70th NYSE percentiles. The following work is just an expansion of the equations reported in Fig.9 associated with the 2x3 sort, which will enhance the ease of comprehension. According to the sorting process being analysed, the four factors augmenting the index model are defined as:

- *Small Minus Big*$_{FF}$

$$SMB\left\{\begin{array}{c} SMB_{(B/M)} = \frac{1}{3}(Small\ High + Small\ Neutral + Small\ Low) \\ -\frac{1}{3}(Big\ High + Big\ Neutral + Big\ Low) \\ SMB_{(OP)} = \frac{1}{3}(Small\ Robust + Small\ Neutral + Small\ Weak) \\ -\frac{1}{3}(Big\ Robust + Big\ Neutral + Big\ Weak) \\ SMB_{(INV)} = \frac{1}{3}(Small\ Conservative + Small\ Neutral + Small\ Aggressive) \\ -\frac{1}{3}(Big\ Conservative + Big\ Neutral + Big\ Aggressive) \end{array}\right\} = \frac{1}{3}(SMB_{(B/M)} + SMB_{(OP)} + SMB_{(INV)})$$

- *High Minus Low*$_{FF}$

$$HML = \frac{1}{2}(Small\ High + Big\ High) - \frac{1}{2}(Small\ \ Low + Big\ Low)$$

- *Conservative Minus Aggressive*$_{FF}$

$$CMA = \frac{1}{2}(Small\ Conservative + Big\ Conservative) - \frac{1}{2}(Small\ \ Aggressive + Big\ \ Aggressive)$$

- *Robust Minus Weak*$_{FF}$

$$RMW = \frac{1}{2}(Small\ Robust + Big\ Robust) - \frac{1}{2}(Small\ \ Weak + Big\ \ Weak)$$

Finally, it should be noted that by anyone of the pair of adjectives employed in the factors' definition, we are implying that one is a proper superset of the other. Precisely, by *Small High* we are implying a $\mathcal{A} \supset \mathcal{B}$ relationship where, in this case, $\mathcal{A} = Small$ and $\mathcal{B} = High$.

3.2.2. Model and variable construction

Discussed hereinabove, the model constructed for the purpose of this empirical dissertation could be theoretically described as an extension of the Fama and French 5-factor model, due to the inclusion of three R&D metrics, sorted in pairs to produce three distinct model. Moreover, the descriptive statistics performed on the collected sample suggest that investors may attach different valuation to individual R&D efforts, to the extent as the activity is able to create value in its context of application. For this reason the model will be performed on industry portfolios.
The following model considers Fama and French factors as control variables, and R&D factors as the variables of interest. Just for the purpose of visualization, we report the general form of the model in which a nonspecific factor, namely Research and Development Metrics (RDM), is introduced.

General model

$$R_{it} - R_{Ft} = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + RDM + e_{it}$$

We employ three different metrics to summarise different facets of Research and Development activity. The following are best interpreted in the context of Fama and French factors' definitions, whose construction will be discussed later in this chapter. The proposed Research and Development Metrics (RDM) are:

- RDP [Research & Development Proportion]

  Defined as the proportion of R&D stock (RDS) to total asset both lagged by one year. It addresses the question: *How big is R&D stock compared to the stock of total asset?*

$$RDP_t = \frac{R\&D\ Stock_{t-1}}{Total\ Asset_{t-1}} = \frac{RDS_{t-1}}{Total\ Asset_{t-1}}$$

- RDRG [Research & Development Relative Growth]

  Defined as the change in R&D expenditure (RDE) standardized by the change in total asset. The metric can be equivalently defined as the ratio of investment in R&D to total investment. Moreover, expenditure is intended as a by-product of the activity (amortization), hence employed to proxy R&D efforts. It essentially addresses the question: *How many units did R&D expenses increase/decrease for a unit change in total asset?*

$$RDRG_t = \frac{d(R\&D\ Expenditure_{t-1})}{d(Total\ Asset_{t-1})} = \frac{RDE_{t-1} - RDE_{t-2}}{Total\ Asset_{t-1} - Total\ Asset_{t-2}}$$

- RDRA [Research & Development Relative Acceleration]

  Defined as the growth rate of R&D expenses compared to the growth rate of total asset. The metric computes the growth rate in excess of that of total asset. It addresses the question: *How much faster/slower did R&D expenditure grew/shrink compared to the growth rate of total asset?*

$$RDGA_t = \frac{\frac{RDE_{t-1} - RDE_{t-2}}{R\&D\ Stock_{t-1}}}{\frac{Total\ Asset_{t-1} - Total\ Asset_{t-2}}{Total\ Asset_{t-1}}}$$

After some manipulation the metric becomes:

$$RDGA_t = \frac{RDRG_t}{RDP_t}$$

Having defined the factors to be plugged into the final model, we now illustrate the mathematical definition of R&D Stock (RDS) and R&D Expenditure (RDE), measurements required for the calculation of the aforementioned metrics. In order to compute both measures we refer to the existing literature to identify a proper coefficient of annual amortization of R&D figures. Ballester, Garcia-Ayuso and Livnat (2003) document that investors consider a high proportion of R&D expenses to provide future benefits to the firm, specifically they find that on average 88.2% of current expenses are considered to generate benefits beyond the year of recognition. Hall, Cummins, Laderman and Mundy (1988) report an R&D amortization rate of 15% found by examining a database on Research and Development activity compiled by the National Bureau of Economic Research (NBER). Finally, we adopt the formulations for RDS and RDE suggested by Chan, Lakonishok and Sougiannis (1999) along with a 20% amortization rate. Below we define the two variables accordingly.

- RDS [Research and Development Stock]

  Represents the share of current and past R&D spending that is still considered to be an asset to the entity, subsequent to amortization adjustment. It assumes that amortization is constant, that follows a linear relationship and most importantly that the entire expenditure in R&D counts towards the creation of an economic assets. This assumption implies that R&D project do not lead to unsuccessful results.

$$RDS_{it} = RD_{it} + 0.8 * RD_{it-1} + 0.6 * RD_{it-2} + 0.4 * RD_{it-3} + 0.2 * RD_{it-4}$$

- RDE [Research and Development Expenditure]

  Represents the periodic amortization of the R&D stock capital asset. Because in RDS we imply that all expenditure is counted toward the creation of an asset, the only source of R&D expenditure becomes that of capital amortization. Hence,

$$RDE_{it} = 0.2 * (RD_{it-1} + RD_{it-2} + RD_{it-3} + RD_{it-4} + RD_{it-5})$$

  It should be noted that in both formulas, RD_{it} is the spending in Research and Development of company i at time t.

Following the estimation of RDS and RDE, we compute the values of the three metrics for every firm in the sample and calculate the arithmetic average of their values for every $t = 1, 2, 3, \dots, 74$. Using the average estimates of RDP, RDRG and RDRA, we construct three distinct models featuring RDP x RDRG, RDP x RDRA and RDRG x RDRA, respectively. Below we report the three specifications.

Model 1

$$R_{it} - R_{Ft} = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + p_i RDP_t + g_i RDRG_t + e_{it}$$

Model 2

$$R_{it} - R_{Ft} = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + p_i RDP_t + q_i RDRA_t + e_{it}$$

Model 3

$$R_{it} - R_{Ft} = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + g_i RDRG_t + q_i RDRA_t + e_{it}$$

For any one of the specifications, the control factors and the regressand are defined in the same way as for the original 5-factor model with the only exception being that of the subscripts. In the model, i will obviously refer to different portfolios than those examined in Fama and French (2014). Nevertheless, despite the equivalent theoretical definitions, we opt for a simpler construction methodology being conscious of its practical implications. The practical structuring of the model is conducted using the R programming language and different specific packages collectively known as *Tidyverse*.

The general model regresses portfolio's gross returns on systematic, non-systematic and R&D related factors. It ignores financial market frictions such as transaction costs and taxation.

Industry portfolios are equally weighted portfolios constructed by matching firms' ticker symbols with the respective Global Industry Classification Standard (GICS) code. In this way we are able to sort firms according to industry and in this way compute portfolios' average returns. Many studies adopting a similar industry approach make use of Standard Industrial Classification (SIC) codes in order to categorize individual stock. We instead opt for a GICS assisted process following Kile and Phillips (2009)*'s* documentation of GICS's improved efficiency when considering High-Tech firms. Even though this study is not exclusively focused on such sector, it still makes sense to adopt the GICS classification, in that descriptive statistics and common business knowledge suggest that High-Tech is one of the most heavily invested industries in R&D activities.

Each specification is estimated on a sample comprising all firms in the Russell 3000 index. Differently from Fama and French (2014), the model is built on quarterly observation. For this reason the risk free assets is identified in the constant maturity 3-month Treasury bill. Being a short period interest rate, the quarterly risk free rate is obtained from the quoted APR using the simple rule. For each firm in the index historical returns are computed as

$$R_{it} = \frac{Market\ capitalization_{it}}{Market\ capitalization_{it-1}} - 1 \tag{4}$$

An identical process is used to estimate market index return. Being a value-weighted index, we use the return on Russell 3000 as a proxy for the return of the CAPM theoretical market portfolio. At this stage, the regressand and the market factor can be estimated by a simple subtraction whose form is reported in both the general model and in its specifications.

From this point onward, even though we use the same ranking variables, we adopt a different sorting approach than that used by Fama and French (2014) by ignoring the initial size filtering.

- *Small Minus Big*

The entire sample is sorted into Small Capitalization and Big Capitalization stocks according to Russell's size breakpoints. These are the lowest 0.75 and highest 0.05 percentiles of the market capitalization ordered sample, respectively.

Index	Companies included (based on descending total market capitalization)
Russell 3000E Index	Companies #1–4,000 or 100% of the eligible securities
Russell 3000 Index	Companies #1–3,000
Russell Top 50 Mega Cap Index	Companies #1-50
Russell Top 200 Index	Companies #1–200
Russell Top 500 Index	Companies # 1-500
Russell 1000 Index	Companies #1–1,000
Russell Midcap Index	Companies #201–1,000
Russell 2000 Index	Companies #1,001–3,000
Russell 2500 Index	Companies #501–3,000
Russell Microcap Index	Companies #2,001–4,000

Fig.10: Russell composition of several distinct indexes according to market capitalization.[15] *Source: FTSE Russell.*

We calculate the reported breakpoints according to *Fig.9*. The logic is that, because Russell defines its Midcap index as comprising firms found between the 201st and 1000th position on the order set of all 4000 eligible firms, it follows that Small capitalization are found in the bottom 3000 firms and Big capitalization stock are found in the top 200 firms. This implies size breakpoints of 200/4000 and 3000/4000, or equivalently 0.05 and 0.75. Because we are not using the entire set of eligible firms, we apply the just calculated percentiles to the available sample based on Russell 3000 composition. The sorting process yields *Small* and *Big* market cap portfolios. Historical market capitalization of individual firms in any one of the constructed portfolios are matched by ticker symbol and by date in order to compute quarterly returns for each $t = 1, 2, 3, \dots, 74$, as in (4). In this way we obtain approximately $74 \times \left[|\overline{Small}| + |\overline{Big}| \right]$ estimates of historical returns, where $|x|$ denotes the magnitude of a set, or equivalently, the number of elements in a specific portfolios. Arithmetic average returns are calculated for both portfolios and for each t. The last process yields $\bar{r}_{SMALL_t}$ and $\bar{r}_{BIG_t}$. The final factor is constructed as:

$$SMB_t = SMALL_t - BIG_t \qquad \text{or, equivalently} \qquad \bar{r}_{SMB_t} = \bar{r}_{SMALL_t} - \bar{r}_{BIG_t}$$

The approach used for computing average returns is true for all Fama and French factors.

---

15 FTSE Russell | Russell U.S. Equity Indexes Construction and Methodology, v4.0, August 2019

- *High Minus Low*

The entire sample of firms is partitioned into High Book to Market firms and Low Book to Market Firm. In order to sort firms into different portfolios we adopt the same breakpoints suggested by Fama, French (2014) in their 2x3 sort. These are the 30th and 70th percentiles. We construct Book Value of Equity as:

$$Book\ value\ of\ equity_{it} = Total\ Assets_{it} - Total\ Liabilities_{it}$$

Successively, the Book to Market ratio is computed as:

$$\frac{B_{it}}{M_{it}} = \frac{Book\ value\ of\ equity_{it-1}}{Market\ capitalization_{it-1}}$$

Finally, we order the sample of firms according to decreasing values of the ratio. Applying the aforementioned percentiles produces two portfolios named as *High* and *Low*. The process ultimately yields the following returns.

$$\bar{r}_{HML_t} = \bar{r}_{HIGH_t} - \bar{r}_{LOW_t}$$

The sorting process involved in the construction of both *Robust Minus Weak* and *Conservative Minus Aggressive* factors makes use of these same breakpoints values.

- *Robust Minus Weak*

The sample of firms is ordered according to operating profitability. This measure is defined and computed according to the following formula.

$$Operating\ profitability_{it} = \frac{Operating\ profit_{it-1} - Interest\ expenses_{it-1}}{Book\ value\ of\ equity_{it-1}}$$

Following the formation of *Robust* and *Weak* and the computation on average returns we obtain the following:

$$\bar{r}_{RMW_t} = \bar{r}_{ROBUST_t} - \bar{r}_{WEAK_t}$$

- *Conservative Minus Aggressive*

The sample of firms is ordered according to investment intensity. The ranking variable is defined and computed according to the following formula.

$$Investment\ intensity_{it} = \frac{Total\ Asset_{t-1} - Total\ Asset_{t-2}}{Total\ Asset_{t-2}}$$

Ranking the sample according to Investment intensity we obtain *Conservative* and *Aggressive* portfolios. The usual process yields the following:

$$\bar{r}_{CMA_t} = \bar{r}_{CONSERVATIVE_t} - \bar{r}_{AGGRESSIVE_t}$$

According to Fama and French (2014) the growth of Book Value of Equity and the growth in Total Assets lead to identical results.

Finally, we regress industry's portfolio excess returns on the estimates of $R_{Mt} - R_{Ft}$, SMB_t, HML_t, RMW_t, CMA_t, RDP_t, $RDRG_t$ and $RDRA_t$. The analysis yields *Model 1*, *Model 2* and *Model 3* for each industry group, for a total of seventy-two specifications.

We display and discuss the results in following sub-chapter.

3.3 Results and implications

The *general model* makes use of seven independent variables for each of the three specifications. Two of these variables are employed in order to capture the effect of R&D investments on industries' excess returns. The remaining five factors are control variables taken from Fama and French (2014). When firm's tickers are matched with GICS codes, the data sample is sorted into 24 portfolios, one for each industry. Each portfolio is then regressed on the three combinations of R&D metrics. Hence, a total of seventy-two models are estimated. The entire framework of analysis use a significance level of $\alpha = 0.05$ and uses OLS techniques to estimate factor loadings. Accordingly, factors whose p-value lies between:

- $0.05 \geq p - value > 0.01$, are referred to as *significant;*
- $0.01 \geq p - value > 0.001$, as *very significant*;
- $0.001 \geq p - value$, as extremely significant.

Furthermore, the market index excess return factor is named as ERM in the models' results. Finally, we are going to report and critically evaluate results in an ordered way without the use of tables. For this reason, we are only going to report results of industries that show statistically significant results on R&D related factors. These are the proportion of Research and Development to Total Asset, the ratio of Research and Development Expenditure Growth to the Growth of Total Asset and Acceleration of total Research and Development efforts. Tabular results for all seventy-two models are placed in Appendix A. We now proceed with results discussion.

With respect to the ***Energy (1010)*** sector, we observe that only *Model 1* yields an extremely significant coefficient on RDRG of 2.01. The sector also shows four statistically significant control variables while ERM is never significant for the all specifications. In *Model 1* we observe very significant coefficients on SMB and CMA of -0.86 and 1.84 respectively, and extremely significant coefficients on HML and RMW of -3.76 and -2.73. In *Model 2* we observe very similar results except for the coefficient on CMA being only significant and estimated at 1.16 In *Model 3*, the four control variable are all associated with a p-value smaller than 0.001. All three models seem to capture much of the variation of Energy's excess returns reporting associated R^2 of 0.96, 0.96 and 0.84, respectively. In ***Materials (1510)*** we observe that the intercept, ERM and SMB are always associated to statistically significant coefficients. Among the three factors only ERM is always extremely significant. *Model 1* yields a very significant coefficient of -0.05 on RDRG. *Model 3* shows very similar results but the coefficient on RDRG is now associated with a greater p-value. All three models show R^2 values between 0.40 and 0.46. The three models on ***Capital Goods (2010)*** yields extremely significant coefficient on both the intercept and ERM. *Model 1* and *Model 2* both report a significant coefficient on RDP of -0.15. Overall, the two models have a low to moderate explanatory power in that their R^2 are 0.42 and 0.41 respectively. For the ***Commercial and professional services (2020)*** sector we observe that all specifications estimate very significant intercept coefficients. These are all in the hundredth unit. They also report extremely significant coefficients on ERM of 0.70 for the first model and 0.73 for the remaining two models. Only *Model 1* reports a very significant coefficient of -0.03 on RDRG and a moderate explanatory power with an R^2 of 0.51. The ***Consumer services (2530)*** industry reports extremely significant coefficient on the intercept and ERM for all three models. In *Model 1* the RDRG factor is associated with a significant coefficient of -0.04. We observe that the specifications can predict only 42% of the industry's returns total variance. With respect to ***Households and Personal products (3030)*** we observe well-estimated control variables. In *Model 1* and *Model 2* the coefficient on SMB, RMW and RMW are extremely significant and have almost equal magnitudes. In *Model 3* the coefficient on CMA is also extremely significant. Its coefficients' magnitudes are more than double that of the previous two models except for HML which is approximately three times less. *Model 1* reports a very significant coefficient on RDRG of 0.55 and an extremely high explanatory power with an R^2 of 0.97. The general model yields the best specifications results in terms of statistical significance for ***Banks (4010)***. *Model 3* yields extremely significant coefficients on SMB, HML and CMA. *Model 1* and *Model 2* report exactly the same coefficient and associated standard errors. We observe extremely significant coefficient on the intercept, SMB, HML, RMW and CMA while the coefficient on ERM is only significant. RDP is associated a very significant factor loading of -0.42. The first two models are able to capture almost the entire variance of Bank's returns with an R^2 of 0.93. ***Technology Hardware and equipment (4520)*** yields extremely significant coefficient on ERM and on RDRG in *Model 1*. We observe a very high negative coefficient on RDRG (-3.19). *Model 2* does not show any significance at all. In *Model 3*, ERM is significant with an extreme coefficient of 6.39 and RDRG is extremely significant with a much lower coefficient (-1.28) with respect to that in *Model 1*. The R^2 varies from one extreme to the other taking on the values of 0.98, 0.05 and 0.42. For the ***Semiconductors and Semiconductors equipment (4530)*** industry we observe very high and extreme magnitudes of control coefficient, usually hovering around [-1,1]. *Model 1* and *Model 2* report extremely significant coefficient on SMB, HML, RMW and CMA. *Model 1* also reports a significant intercept coefficient and an extremely significant factor loading on RDRG of 4.20. From *Model 3* we can observe an extremely significant coefficient on CMA and a very significant coefficient with value 2.09 on RDRG. All three specifications report R^2 of 1 meaning that the variance of industry returns is entirely captured by the independent factors. *Model 1* and *Model 2* on ***Utilities (5510)*** yield the exact same results. Both report very significant intercept coefficients and extremely significant factor loadings on ERM. The coefficient (-0.10) on RDP is significant. *Model 3* instead reports equivalent results on intercept and ERM but no significance of

R&D metrics. All three models have low explanatory power with R^2 values of 0.34, 0.33 and 0.29 respectively. ***Real Estate (6010)*** provides unique results in that, *Model 1* estimates a very significant coefficient on RDP of -0.25 and an extremely significant coefficient on RDRG of -0.11. Extremely significant values are also observed for the intercept and ERM for both *Model 1* and *Model 2*. The latter has a very significant coefficient on RDP of similar magnitudes to that of *Model 1*. *Model 3* reports a very significant intercept coefficient, an extremely significant loading on ERM and a significant coefficient on RDRG of -0.04. *Model 1* exhibit moderate explanatory power with an R^2 of 0.52 while *Model 2* and *Model 3* both exhibit R^2 values of 0.29.

Thirteen out of the twenty-four industries do not show any significant coefficient on RDP, RDRG and RDRA. These are Transportation (2030), Automobiles and components (2510), Consumer Durables and Apparel (2520), Retailing (2550), Food & staples retailing (3010), Food, Beverage and Tobacco (3020), Health care equipment and services (3510), Pharmaceuticals Biotechnology and Life Sciences (3520), Diversifies financials (4020), Insurance (4030), Software and services (4510), Telecommunication services (5010), Media and entertainment (5020).

Exhibit 2. R&D Metrics for Industries with at least one significant coefficient.

	Metrics	Model 1	Model 2	Model 3
Energy (1010)	RDP	-2.87	-2.56	
	RDRG	2.01***		0.44
	RDRA		-0.01	-0.01
Materials (1510)	RDP	-0.05	-0.06	
	RDRG	-0.05**		-0.00
	RDRA		0.00	0.00
Capital goods (2010)	RDP	-0.15*	-0.15*	
	RDRG	-0.02		-0.01
	RDRA		0.00	0.00
Commercial and prof. services (2020)	RDP	-0.10	-0.11	
	RDRG	-0.03**		-0.01
	RDRA		-0.00	-0.00
Consumer services (2530)	RDP	-0.10	-0.11	
	RDRG	-0.04*		-0.01
	RDRA		-0.00	-0.00
Household & Personal products (3030)	RDP	0.31	0.41	
	RDRG	0.55**	-0.00	0.14
	RDRA			-0.00
Banks (4010)	RDP	-0.42**	-0.42**	
	RDRG	0.02		0.03
	RDRA		-0.00	-0.00
Technology Hardware and equipment (4520)	RDP	-0.12	-0.72	
	RDRG	-3.19***		-1.28***
	RDRA		0.00	0.00
Semiconductors and Semiconductors Equip (4530)	RDP	-4.89	-4.21	
	RDRG	4.20***		2.09**
	RDRA		-0.03	-0.02
Utilities (5510)	RDP	-0.10*	-0.10*	
	RDRG	-0.01		-0.01
	RDRA		0.00	0.00
Real Estate (6010)	RDP	-0.25**	-0.28**	
	RDRG	-0.11***		-0.04*
	RDRA		-0.00	-0.00

Source: Author

Hence, from the results reported in Exhibit 2. we may deduce that R&D efforts are positively valued when performed in the Semiconductors and Semiconductors equipment, Energy and Household and personal product services. The effect is apparently opposite in the Technology Hardware and equipment, Banks, Capital Goods and Real Estate industries. These are industries for which the models estimate high enough negative coefficients. The remaining four industries for which the model report statistically significant coefficient are Materials, Commercial and professional services, Consumer services and Utilities. We mentioned them separately because even if they do show negative and statistically significant coefficient, their magnitudes are extremely

small. Since we employed ratios rather than absolute values, coefficients on the hundredth scale imply that a consistent effect is only perceivable when the underlying factors experience unrealistic variations.
According to results, we argue that for the first industry, Model 1 implies that, keeping every other factor fixed, one unit change in the value of RDRG is, on average, amplified by a factor of four into industry's excess returns. The same reasoning can be applied to all the other estimated coefficients. A final observation is that RDRA is never significant and hardly ever different from zero therefore it could be excluded from the specifications or tested as a standalone.

Although this framework of analysis yields several statistically significant coefficients, we must be wary of the produced results and the implications that can be drawn from them. Several problems are indeed linked to the approach used in constructing the model.
A first problem is posed by serial correlation. It refers to the situation in which observation of the same variable are correlated between dates t. Therefore, if an observation in $t-1$ is able to predict the successive one, the estimated coefficient would show irrelevant numbers that do not truly reflect the extent of their explanatory power on the dependent variable but rather a value that results from a combination of the real effect and of the effect from autocorrelation.

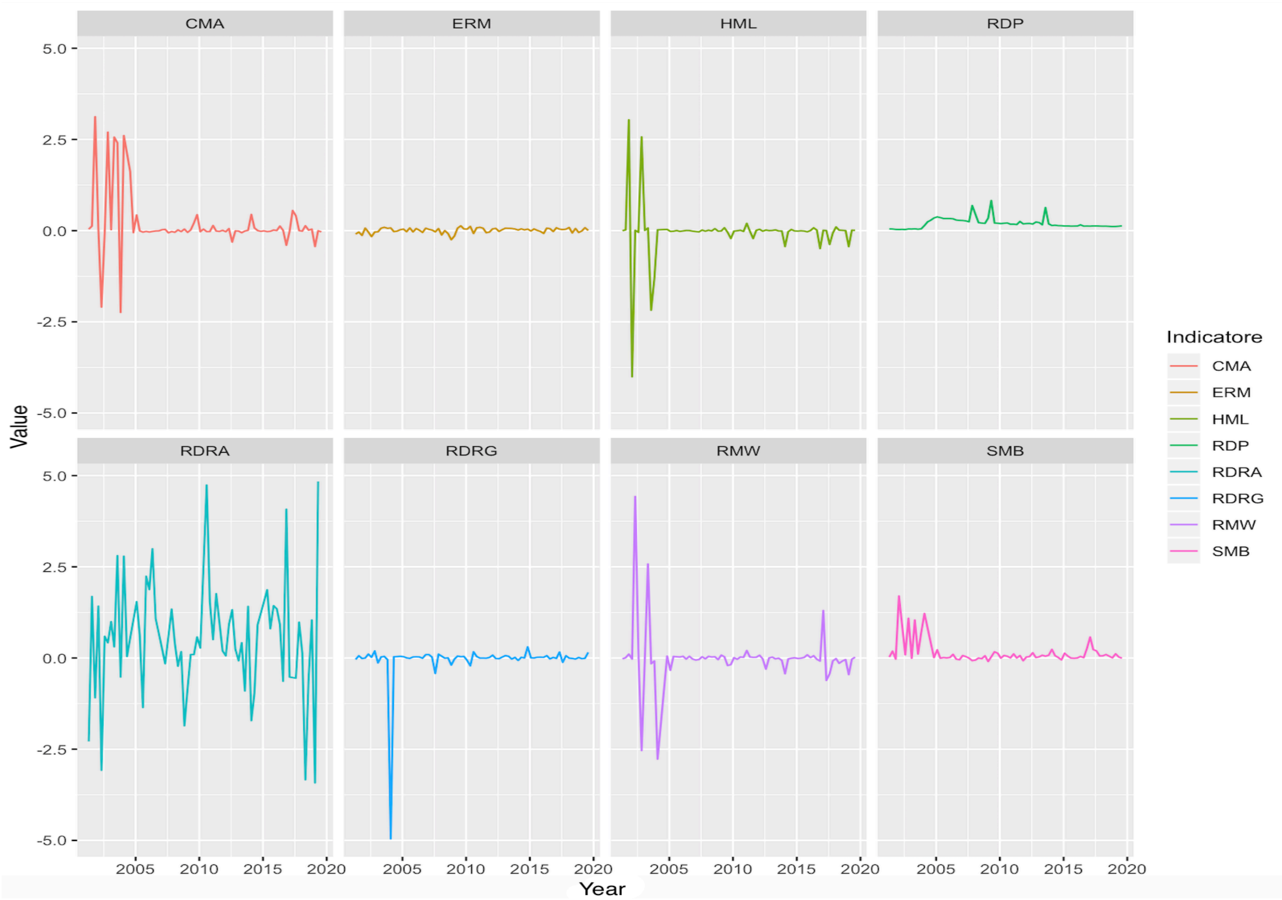

Fig.11: Factors' variability. Source: Author.

Looking at factors' distributions over time can usually help to spot the bias. Factors that exhibit trends over time may in fact be affected by autocorrelation. In this case CMA, HML, RDRA, RMW and SMB do apparently show high enough variability to reassure of their independence in time. Yet, Fig.11 only serves figurative purposes in providing a critical view on estimated results so we prefer to capture the entire variation of factors such as RDRA at the expense of the remaining ones. These would need to be observed at a different value scale in order to have a clear picture of their distribution over time.

A second problem is posed by multicollinearity, being a situation in which independent variables are correlated. When the problem is severe enough it leads to meaningless results. It violates the purpose of regression analysis itself. Estimated coefficients are interpreted as the mean change in the dependent variable for each unit change in an independent when holding all other variables fixed. Therefore, if two independent variables are correlated it becomes difficult to interpret the explanatory role of a variable since a unit change in the first prompts a change of the second one. In addition, the problem can be either structural, model related, or data multicollinearity, already embedded in the data. For this reason we briefly examine the correlation coefficient of independent variables. Correlation tables for individual industry's excess return and independent variables are provided in Appendix 2.

Exhibit 3 ~ Factors' correlation

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
CMA	1.000	-0.065	0.979	0.029	-0.022	-0.002	-0.988	0.957
ERM	-0.065	1.000	-0.071	0.145	0.104	-0.084	0.068	-0.036
HML	0.979	-0.071	1.000	0.025	-0.020	0.015	-0.942	0.902
RDP	0.029	0.145	0.025	1.000	-0.036	0.051	-0.051	0.065
RDRA	-0.022	0.104	-0.020	-0.036	1.000	0.039	0.022	-0.011
RDRG	-0.002	-0.084	0.015	0.051	0.039	1.000	0.004	-0.010
RMW	-0.988	0.068	-0.942	-0.051	0.022	0.004	1.000	-0.974
SMB	0.957	-0.036	0.902	0.065	-0.011	-0.010	-0.974	1.000

Source: Author

Exhibit 3 presents correlation coefficient among employed factor. We do not highlight all relevant coefficients, as the table is symmetric about the diagonal. In yellow we can observe extremely high correlation coefficients among all control variables except for the factor representing market index excess return. These coefficients are in fact equivalent to both negative and positive perfect correlation. Being control variables such high correlation coefficient should not, in principal, threat the validity of the produced results. High R^2 values may partly be a product of the observed level of correlation. Furthermore, it comes at a surprise to observe that RDP, RDRA and RDRG are virtually uncorrelated among each other. The following, because the mathematical definition of the factor presents itself both data and structural links.

Chapter 4

Conclusion

The thesis aims at exploring the relationship between Research & Development expenditure and excess market returns. In doing so we use an industry-oriented approach, being conscious that such activity is not valued in isolation but in a context aware manner. We collect and provide both general and specific information in order to characterize Research & Development as an intangible asset in itself and as an element part of a wider environment.

We develop three metrics in order to capture different aspects of Research & Development measurement, namely Proportion, Growth and Acceleration. We construct these variables by adopting a model of Research and Development stock and expense formation from Chan, Lakonishok and Sougiannis (1999). Attempting to remove estimation noise from the variable of interest we complement the metrics with a simplified version of the Fama and French 5-factor model. We opt for a different construction of control variables as they did not play a central role in the study and the only objective was that of capturing the variation of industry excess returns attributable to systematic and non-systematic factors.

The model produces interesting results. Out of the twenty-four industries, Research and Development yields statistically significant coefficients in eleven of those. In particular, the model suggests that financial markets attach positive and considerable value the proportion of Research and Development to Total Asset in the Energy, Semiconductors and Household and personal products industries. At the same time the model suggests an opposite effect for Technology Hardware, Banks and Real Estate industries.

This study represents a starting point for personal future research on the world of intangible assets. Numerous approaches can be employed to test and explore just as many individual facets of Research and Development. This dissertation may have produced more goals than recommendations with respect to my research activity. For this reason it is my intention to further explore this knowledge area using different sampling approaches and more advanced techniques such as Principal Component Analysis in order to better uncover the underlying relationship.

Bibliography

- Aharoni, G.,Grundy,B.,Zeng,Q., 2013. Stock returns and the Miller Modigliani valuation formula: revisiting the Fama French analysis. Journal of Financial Economics 110, 347–357.
- Alex Coad & Rekha Rao, 2006. "Innovation and Firm Growth in High-Tech Sectors: A Quantile Regression Approach," LEM Papers Series 2006/18, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.
- Arthur, W. B. (2009). *The nature of technology: what it is and how it evolves*. New York: Free Press.
- Baker, N., Freeland, J., 1975. Recent advances in R&D benefit measurement and project selection methods. Manag. Sci. 21 (2), 1164–1175. doi: 10.1287/mnsc. 21.10.1164 .
- Brand Finance Institute. (2018). GIFT 2018
- Bronwyn H. Hall, 1998. "Innovation and Market Value" NBER Working Papers 6984, National Bureau of Economic Research, Inc.
- Bublitz, B. and Ettredge, M. (1989) "The information in discretionary outlays: advertising, research and development", Accounting Review, 64: 108–24.
- Corrado, C., Haltiwanger, J. C., & Sichel, D. E. (2005). *Measuring capital in the new economy*. Chicago: The University of Chicago Press. doi: 10.7208/chicago/9780226116174.001.0001
- Edvinsson, L., Malone, M.S. , March 1997. Intellectual Capital: Realizing Your Company's True Value by Finding its Hidden Brainpower, pp. 26–28 *Collins*
- Fama, E. F., & French, K. R. (2014). *A five-factor asset pricing model*. Elsevier. doi: http://dx.doi.org/10.1016/j.jfineco.2014.10.010
- FTSE Russell | Russell U.S. Equity Indexes Construction and Methodology, v4.0, August 2019
- *Gics ; Global Industry Classification Standard*. (0AD).
- Glaeser, E. L. (2011). *Triumph of the city*. London: Macmillan.
- Hall, Bronwyn H. 1993a. "The Stock Market Valuation of R&D Investment during the 1980s", American Economic Review 83: 259-264
- Hall, Bronwyn H. 1993b. "Industrial Research during the 1980s: Did the Rate of Return Fall?" Brookings Papers on Economic Activity Micro 2: 289-344
- Hall, Bronwyn H., Clint Cummins, Elizabeth S. Laderman, and Joy Mundy, 1988, The R&D master file documentation, NBER Technical Working Paper no. 72.
- Harhoff, Dietmar; Narin, Francis; Scherer, Frederic M.; Vopel, Katrinl (1997) : Citation frequency and the value of patented innovation, WZB Discussion Paper, No. FS IV 97-26, Wissenschaftszentrum Berlin für Sozialforschung (WZB), Berlin

- Haskel, J., & Westlake, S. (2018). *Capitalism without capital: the rise of the intangible economy*. Princeton, NJ: Princeton University Press.
- Hirschey, Mark, Vernon J. Richardson, and Susan Scholz, 1998. "Value relevance of Nonfinancial Information: The case of Patent data", University of Kansas School of Business.
- Hirshleifer, D., Lim, S., Toeh, S.H., 2009. Driven to distraction: extraneous events and under reaction to earnings news. J. Finance 64 (5), 2289–2325. doi: 10. 1111/j.1540-6261.2009.01501.x
- Hsu, D.H., Ziedonis, R.H., 2013. Resources as dual sources of advantage: Implications for valuing entrepreneurial-firm patents. Strategic Management Journal 34 (7), 761–781. doi: 10.1002/smj.2037 .
- J. Doukas and L.N. Switzer, "The Stock Market's View of R&D Spending and Market Concentration," Journal of Economics and Business (May 1992), pp. 95
- Jarrel, G. A., Lehn, K. and Marr, M. W. (1985) Institutional Ownership, Tender Offers and Long Term Investments. Washington, DC: SEC.
- Keith W. Chauvin and Mark Hirschey (1993), Advertising, R&D Expenditures and the Market Value of the Firm *Financial Management* Vol. 22, No. 4 (Winter, 1993), pp. 128-140 DOI: 10.2307/3665583
- Kile C, Phillips M (2009) Using industry classification codes to sample high-technology firms: Analysis and recommendations. J Account Audit Finance 24(1):35–58. doi: 10.1177/0148558X0902400104
- KPMG IFRG Limited. (2017). *Ifrs compared to Us Gaap*. London
- Lev, B. and Sougiannis, T. (1996) 'The capitalization, amortization and value relevance of R&D', Journal of Accounting and Economics, 21: 107–38.
- Lev, B., Sougiannis, T., 1999. Penetrating the book-to-market black box: the R&D effect. J. Bus. Finance Account. 26 (3), 419–449. doi: 10.1111/1468-5957. 00262 .
- Louis K. C. Chan, Josef Lakonishok, Theodore Sougiannis (1999). "The Stock Market Valuation of Research and Development Expenditures" Journal of Finance, American Finance Association, vol. 56(6), pages 2431-2456, December.
- M. Statman and J.F. Sepe, "Project Termination Announcements and the Market Value of the Firm," Financial Management (Winter 1989), pp. 74-81
- Marta Ballester , Manuel Garcia-Ayuso & Joshua Livnat (2003) The economic value of the R&D intangible asset , European Accounting Review, 12:4, 605-633, DOI: 10.1080/0963818031000162843[7]
- National Bureau of Economic Research, Kenneth J. Arrow (1962). Economic Welfare and the Allocation of Resources for Invention. *The Rate and Direction of Inventive Activity: Economic*

and Social Factors (pp. 609–626). Princeton: Princeton University Press. https://doi.org/10.1515/9781400879762-024

- Novy-Marx, R., 2013. The other side of value: The gross profitability premium. Journal of Financial Economics 108, 1–28.
- OECD (2015), *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris, https://doi.org/10.1787/9789264239012-en
- Romer, P. M. (1990). *Endogenous technological change*. Cambridge, MA: NBER. doi: 10.3386/w3210
- S.H. Chan, J.D. Martin, and J.W. Kensinger, "Corporate Research and Development Expenditures and Share Value," Journal of Financial Economics (August 1990), pp.
- Tangible and Intangible Investment in Corporate Finance, 2019. The North American Journal of Economics and Finance 50:100991 DOI: 10.1016/j.najef.2019.100991
- Trajenberg, Manuel. 1990. "A Penny for Your Quotes: Patent Citations and the Value of Innovations" Rand Journal of Economics 21: 172-187.
- United Nations. (2009). *System of national accounts 2008*. New York.
- Woolridge, J. R. (1988) 'Competitive decline and corporate restructuring: is a myopic stock market to blame?', Journal of Applied Corporate Finance, 1: 26–36.
- Yu, Gun Jea & Hong, KiHoon, 2016. "Patents and R&D expenditure in explaining stock price movements" Finance Research Letters, Elsevier, vol. 19(C), pages 197-203.

Appendix A

Model specifications results

Automobiles and components (2510)

	Model 1	Model 2	Model 3
(Intercept)	0.08 *	0.09 *	0.04
	(0.04)	(0.04)	(0.02)
ERM	1.52 ***	1.55 ***	1.47 ***
	(0.32)	(0.32)	(0.31)
SMB	-0.05 **	-0.05 **	-0.05 **
	(0.02)	(0.02)	(0.02)
HML	-0.01	-0.01	-0.02 **
	(0.01)	(0.01)	(0.01)
RMW	-0.02	-0.02	-0.00
	(0.02)	(0.02)	(0.01)
CMA	0.01	0.01	0.04 *
	(0.03)	(0.03)	(0.02)
RDP	-0.20	-0.21	
	(0.16)	(0.16)	
RDRG	-0.01		0.00
	(0.04)		(0.02)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.32	0.33	0.32

Diversified financials (4020)

	Model 1	Model 2	Model 3
(Intercept)	0.23	0.23	0.34
	(0.51)	(0.52)	(0.31)
ERM	2.71	2.49	3.86
	(4.18)	(4.22)	(4.29)
SMB	-1.41 ***	-1.41 ***	-1.64 ***
	(0.23)	(0.23)	(0.23)
HML	-1.18 ***	-1.16 ***	-0.66 ***
	(0.19)	(0.19)	(0.09)
RMW	1.09 ***	1.07 ***	0.41 **
	(0.26)	(0.26)	(0.15)
CMA	2.78 ***	2.73 ***	1.70 ***
	(0.40)	(0.40)	(0.22)
RDP	-0.52	-0.46	
	(2.07)	(2.08)	
RDRG	0.51		0.11
	(0.48)		(0.32)
RDRA		-0.01	-0.01
		(0.03)	(0.03)
N	72	72	74
R2	0.58	0.57	0.50

Insurance (4030)

	Model 1	Model 2	Model 3
(Intercept)	0.05 **	0.05 **	0.03 *
	(0.02)	(0.02)	(0.01)
ERM	0.68 ***	0.69 ***	0.66 ***
	(0.15)	(0.15)	(0.14)
SMB	-0.00	-0.00	-0.01
	(0.01)	(0.01)	(0.01)
HML	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.00)
RMW	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.00)
CMA	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)
RDP	-0.13	-0.13	
	(0.07)	(0.07)	
RDRG	-0.01		-0.00
	(0.02)		(0.01)
RDRA		0.00	0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.29	0.28	0.26

Semiconductors and Semiconductors equipment (4530)

	Model 1	Model 2	Model 3
(Intercept)	1.58 *	1.38	-0.19
	(0.69)	(0.90)	(0.78)
ERM	-2.36	-4.87	-19.12
	(5.69)	(7.33)	(10.70)
SMB	-1.87 ***	-1.85 ***	-0.59
	(0.31)	(0.40)	(0.57)
HML	2.52 ***	2.71 ***	-0.23
	(0.25)	(0.33)	(0.23)
RMW	-4.29 ***	-4.47 ***	-0.62
	(0.35)	(0.45)	(0.37)
CMA	4.33 ***	3.96 ***	10.06 ***
	(0.55)	(0.70)	(0.54)
RDP	-4.89	-4.21	
	(2.81)	(3.62)	
RDRG	4.20 ***		2.09 **
	(0.65)		(0.79)
RDRA		-0.03	-0.02
		(0.05)	(0.08)
N	72	72	74
R2	1.00	1.00	1.00

Banks (4010)

	Model 1	Model 2	Model 3
(Intercept)	0.15 ***	0.15 ***	0.04
	(0.04)	(0.04)	(0.03)
ERM	0.67 *	0.68 *	0.20
	(0.29)	(0.29)	(0.41)
SMB	-0.46 ***	-0.45 ***	-0.41 ***
	(0.02)	(0.02)	(0.02)
HML	-0.06 ***	-0.06 ***	-0.17 ***
	(0.01)	(0.01)	(0.01)
RMW	-0.13 ***	-0.13 ***	0.01
	(0.02)	(0.02)	(0.01)
CMA	0.12 ***	0.12 ***	0.34 ***
	(0.03)	(0.03)	(0.02)
RDP	-0.42 **	-0.42 **	
	(0.14)	(0.14)	
RDRG	0.02		0.03
	(0.03)		(0.03)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.93	0.93	0.85

Energy (1010)

	Model 1	Model 2	Model 3
(Intercept)	0.96	0.87	1.45
	(0.62)	(0.68)	(0.75)
ERM	-0.96	-2.10	6.37
	(5.08)	(5.55)	(10.23)
SMB	-0.86 **	-0.85 **	-2.16 ***
	(0.28)	(0.30)	(0.55)
HML	-3.76 ***	-3.67 ***	-0.87 ***
	(0.23)	(0.25)	(0.22)
RMW	-2.73 ***	-2.82 ***	-6.47 ***
	(0.31)	(0.34)	(0.35)
CMA	1.34 **	1.16 *	-4.60 ***
	(0.49)	(0.53)	(0.52)
RDP	-2.87	-2.56	
	(2.51)	(2.74)	
RDRG	2.01 ***		0.44
	(0.58)		(0.75)
RDRA		-0.02	-0.03
		(0.04)	(0.08)
N	72	72	74
R2	0.96	0.96	0.84

Materials (1510)

	Model 1	Model 2	Model 3
(Intercept)	0.05 *	0.05 *	0.04 **
	(0.02)	(0.02)	(0.01)
ERM	1.00 ***	1.04 ***	1.03 ***
	(0.16)	(0.17)	(0.16)
SMB	-0.02 **	-0.02 *	-0.02 **
	(0.01)	(0.01)	(0.01)
HML	-0.00	-0.00	-0.01 **
	(0.01)	(0.01)	(0.00)
RMW	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)
CMA	0.00	0.00	0.01
	(0.02)	(0.02)	(0.01)
RDP	-0.05	-0.06	
	(0.08)	(0.08)	
RDRG	-0.05 **		-0.03 *
	(0.02)		(0.01)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.46	0.40	0.43

Software and services (4510)

	Model 1	Model 2	Model 3
(Intercept)	1.11	1.13 *	0.43
	(0.56)	(0.56)	(0.33)
ERM	8.03	8.09	7.22
	(4.59)	(4.60)	(4.47)
SMB	0.04	0.04	-0.06
	(0.25)	(0.25)	(0.24)
HML	-0.14	-0.13	-0.02
	(0.21)	(0.20)	(0.10)
RMW	0.12	0.11	-0.02
	(0.28)	(0.28)	(0.15)
CMA	0.24	0.22	0.03
	(0.44)	(0.44)	(0.23)
RDP	-3.78	-3.79	
	(2.27)	(2.27)	
RDRG	0.18		-0.01
	(0.52)		(0.33)
RDRA		-0.01	-0.01
		(0.03)	(0.03)
N	72	72	74
R2	0.09	0.09	0.04

Capital goods (2010)

	Model 1	Model 2	Model 3
(Intercept)	0.07 ***	0.07 ***	0.04 ***
	(0.02)	(0.02)	(0.01)
ERM	0.86 ***	0.87 ***	0.86 ***
	(0.15)	(0.15)	(0.15)
SMB	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)
HML	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.00)
RMW	0.00	0.00	-0.00
	(0.01)	(0.01)	(0.01)
CMA	0.02	0.02	0.01
	(0.01)	(0.01)	(0.01)
RDP	-0.15 *	-0.15 *	
	(0.07)	(0.07)	
RDRG	-0.02		-0.01
	(0.02)		(0.01)
RDRA		0.00	0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.42	0.41	0.36

Food & staples retailing (3010)

	Model 1	Model 2	Model 3
(Intercept)	0.00	0.01	0.04 *
	(0.03)	(0.03)	(0.01)
ERM	0.45 *	0.48 *	0.54 **
	(0.21)	(0.21)	(0.20)
SMB	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)
HML	0.00	0.00	0.00
	(0.01)	(0.01)	(0.00)
RMW	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)
CMA	0.00	-0.00	-0.00
	(0.02)	(0.02)	(0.01)
RDP	0.16	0.15	
	(0.10)	(0.10)	
RDRG	-0.00		0.01
	(0.02)		(0.01)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.12	0.14	0.13

Media & Entertainment (5020)

	Model 1	Model 2	Model 3
(Intercept)	0.08 **	0.08 **	0.04 **
	(0.03)	(0.03)	(0.02)
ERM	1.25 ***	1.28 ***	1.21 ***
	(0.21)	(0.21)	(0.21)
SMB	-0.02	-0.02	-0.02
	(0.01)	(0.01)	(0.01)
HML	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.00)
RMW	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.01)
CMA	0.01	0.01	0.01
	(0.02)	(0.02)	(0.01)
RDP	-0.19	-0.20	
	(0.11)	(0.11)	
RDRG	-0.01		-0.01
	(0.02)		(0.02)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.37	0.38	0.35

Technology Hardware and equipment (4520)

	Model 1	Model 2	Model 3
(Intercept)	0.09	0.29	-0.01
	(0.06)	(0.44)	(0.19)
ERM	2.07 ***	4.27	6.39 *
	(0.53)	(3.56)	(2.63)
SMB	-0.04	-0.04	-0.11
	(0.03)	(0.19)	(0.14)
HML	-0.02	-0.16	-0.04
	(0.02)	(0.16)	(0.06)
RMW	-0.00	0.13	-0.02
	(0.03)	(0.22)	(0.09)
CMA	0.04	0.30	0.07
	(0.05)	(0.34)	(0.13)
RDP	-0.12	-0.72	
	(0.26)	(1.75)	
RDRG	-3.19 ***		-1.28 ***
	(0.06)		(0.19)
RDRA		0.00	0.00
		(0.03)	(0.02)
N	72	72	74
R2	0.98	0.05	0.42

Commercial and professional services (2020)

	Model 1	Model 2	Model 3
(Intercept)	0.04 **	0.05 **	0.03 **
	(0.01)	(0.01)	(0.01)
ERM	0.70 ***	0.73 ***	0.73 ***
	(0.11)	(0.11)	(0.11)
SMB	-0.01	-0.01	-0.01 *
	(0.01)	(0.01)	(0.01)
HML	-0.01	-0.01	-0.01 *
	(0.00)	(0.01)	(0.00)
RMW	-0.00	0.00	-0.00
	(0.01)	(0.01)	(0.00)
CMA	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)
RDP	-0.10	-0.11	
	(0.05)	(0.06)	
RDRG	-0.03 **		-0.01
	(0.01)		(0.01)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.51	0.45	0.44

Food, Beverage and Tobacco (3020)

	Model 1	Model 2	Model 3
(Intercept)	0.06 ***	0.06 ***	0.04 ***
	(0.01)	(0.01)	(0.01)
ERM	0.60 ***	0.63 ***	0.60 ***
	(0.09)	(0.09)	(0.09)
SMB	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.00)
HML	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)
RMW	-0.01	-0.01	-0.00
	(0.01)	(0.01)	(0.00)
CMA	-0.00	-0.00	0.00
	(0.01)	(0.01)	(0.00)
RDP	-0.09	-0.10 *	
	(0.05)	(0.05)	
RDRG	-0.01		-0.00
	(0.01)		(0.01)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.42	0.44	0.41

Pharmaceuticals, Biotechnology and Life Sciences (3520)

	Model 1	Model 2	Model 3
(Intercept)	0.30 **	0.31 **	0.20 ***
	(0.09)	(0.10)	(0.05)
ERM	2.10 **	2.19 **	2.01 **
	(0.78)	(0.78)	(0.75)
SMB	-0.00	-0.00	0.00
	(0.04)	(0.04)	(0.04)
HML	0.03	0.03	0.00
	(0.03)	(0.03)	(0.02)
RMW	-0.04	-0.04	0.00
	(0.05)	(0.05)	(0.03)
CMA	0.00	0.01	0.07
	(0.07)	(0.07)	(0.04)
RDP	-0.46	-0.48	
	(0.38)	(0.39)	
RDRG	-0.09		-0.03
	(0.09)		(0.06)
RDRA		-0.00	-0.00
		(0.01)	(0.01)
N	72	72	74
R2	0.85	0.85	0.85

Telecommunication services (5010)

	Model 1	Model 2	Model 3
(Intercept)	-0.16	-0.17	0.24
	(0.39)	(0.40)	(0.23)
ERM	1.81	1.77	2.60
	(3.23)	(3.24)	(3.10)
SMB	-0.01	-0.01	-0.03
	(0.18)	(0.18)	(0.17)
HML	-0.11	-0.11	-0.01
	(0.14)	(0.14)	(0.07)
RMW	0.16	0.16	0.03
	(0.20)	(0.20)	(0.11)
CMA	0.28	0.27	0.05
	(0.31)	(0.31)	(0.16)
RDP	1.89	1.90	
	(1.60)	(1.60)	
RDRG	0.07		0.02
	(0.37)		(0.23)
RDRA		-0.00	-0.00
		(0.02)	(0.02)
N	72	72	74
R2	0.04	0.04	0.01

Consumer Durables and Apparel (2520)

	Model 1	Model 2	Model 3
(Intercept)	0.34	0.35	0.21
	(0.31)	(0.31)	(0.18)
ERM	3.19	3.28	3.18
	(2.53)	(2.54)	(2.41)
SMB	0.09	0.10	0.06
	(0.14)	(0.14)	(0.13)
HML	-0.03	-0.03	0.03
	(0.11)	(0.11)	(0.05)
RMW	0.13	0.12	0.06
	(0.16)	(0.16)	(0.08)
CMA	0.12	0.11	0.01
	(0.24)	(0.24)	(0.12)
RDP	-0.80	-0.83	
	(1.25)	(1.25)	
RDRG	0.06		0.01
	(0.29)		(0.18)
RDRA		-0.01	-0.01
		(0.02)	(0.02)
N	72	72	74
R2	0.05	0.05	0.04

Health care equipment and services (3510)

	Model 1	Model 2	Model 3
(Intercept)	0.08 ***	0.08 ***	0.06 ***
	(0.02)	(0.02)	(0.01)
ERM	0.63 ***	0.64 ***	0.64 ***
	(0.15)	(0.15)	(0.14)
SMB	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.01)
HML	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.00)
RMW	0.00	0.00	0.00
	(0.01)	(0.01)	(0.00)
CMA	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)
RDP	-0.10	-0.10	
	(0.07)	(0.07)	
RDRG	-0.01		-0.00
	(0.02)		(0.01)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.26	0.25	0.25

Real Estate (6010)

	Model 1	Model 2	Model 3
(Intercept)	0.10 ***	0.11 ***	0.04 **
	(0.02)	(0.02)	(0.01)
ERM	0.80 ***	0.89 ***	0.88 ***
	(0.16)	(0.20)	(0.19)
SMB	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)
HML	0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.00)
RMW	-0.00	-0.00	0.00
	(0.01)	(0.01)	(0.01)
CMA	-0.00	0.00	0.01
	(0.02)	(0.02)	(0.01)
RDP	-0.25 **	-0.28 **	
	(0.08)	(0.10)	
RDRG	-0.11 ***		-0.04 *
	(0.02)		(0.01)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.52	0.29	0.29

Transportation (2030)

	Model 1	Model 2	Model 3
(Intercept)	0.13 *	0.12 *	0.08 *
	(0.05)	(0.05)	(0.03)
ERM	1.05 *	1.01 *	0.95 *
	(0.44)	(0.43)	(0.41)
SMB	-0.01	-0.01	-0.01
	(0.02)	(0.02)	(0.02)
HML	-0.00	-0.00	-0.00
	(0.02)	(0.02)	(0.01)
RMW	-0.01	-0.01	-0.01
	(0.03)	(0.03)	(0.01)
CMA	-0.00	0.00	0.00
	(0.04)	(0.04)	(0.02)
RDP	-0.26	-0.25	
	(0.22)	(0.21)	
RDRG	-0.00		0.00
	(0.05)		(0.03)
RDRA		0.00	0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.10	0.11	0.10

Consumer services (2530)

	Model 1	Model 2	Model 3
(Intercept)	0.06 ***	0.07 ***	0.04 ***
	(0.02)	(0.02)	(0.01)
ERM	0.72 ***	0.75 ***	0.77 ***
	(0.13)	(0.14)	(0.14)
SMB	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)
HML	-0.01	-0.01	-0.00
	(0.01)	(0.01)	(0.00)
RMW	0.00	0.00	-0.00
	(0.01)	(0.01)	(0.00)
CMA	0.01	0.02	0.01
	(0.01)	(0.01)	(0.01)
RDP	-0.10	-0.11	
	(0.07)	(0.07)	
RDRG	-0.04 *		-0.01
	(0.02)		(0.01)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.42	0.37	0.36

Household & Personal products (3030)

	Model 1	Model 2	Model 3
(Intercept)	-0.02	-0.05	0.35
	(0.19)	(0.20)	(0.22)
ERM	-2.36	-2.73	0.21
	(1.53)	(1.65)	(3.00)
SMB	-0.43 ***	-0.43 ***	-0.81 ***
	(0.08)	(0.09)	(0.16)
HML	-1.19 ***	-1.16 ***	-0.32 ***
	(0.07)	(0.07)	(0.06)
RMW	-1.07 ***	-1.09 ***	-2.19 ***
	(0.09)	(0.10)	(0.10)
CMA	0.27	0.23	-1.51 ***
	(0.15)	(0.16)	(0.15)
RDP	0.31	0.41	
	(0.76)	(0.82)	
RDRG	0.55 **		0.14
	(0.17)		(0.22)
RDRA		-0.00	-0.00
		(0.01)	(0.02)
N	72	72	74
R2	0.97	0.97	0.87

Retailing (2550)

	Model 1	Model 2	Model 3
(Intercept)	0.02	0.03	0.03 **
	(0.02)	(0.02)	(0.01)
ERM	1.04 ***	1.06 ***	1.09 ***
	(0.17)	(0.16)	(0.16)
SMB	-0.02	-0.02	-0.02 *
	(0.01)	(0.01)	(0.01)
HML	-0.02 *	-0.02 *	-0.01 *
	(0.01)	(0.01)	(0.00)
RMW	0.01	0.01	0.00
	(0.01)	(0.01)	(0.01)
CMA	0.04 *	0.04 *	0.02 *
	(0.02)	(0.02)	(0.01)
RDP	0.02	0.02	
	(0.08)	(0.08)	
RDRG	0.02		0.01
	(0.02)		(0.01)
RDRA		-0.00	-0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.43	0.45	0.46

Utilities (5510)

	Model 1	Model 2	Model 3
(Intercept)	0.04 **	0.04 **	0.02 *
	(0.01)	(0.01)	(0.01)
ERM	0.46 ***	0.47 ***	0.44 ***
	(0.10)	(0.10)	(0.10)
SMB	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)
HML	-0.01	-0.01	-0.00
	(0.00)	(0.00)	(0.00)
RMW	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.00)
CMA	0.01	0.01	0.00
	(0.01)	(0.01)	(0.00)
RDP	-0.10 *	-0.10 *	
	(0.05)	(0.05)	
RDRG	-0.01		-0.01
	(0.01)		(0.01)
RDRA		0.00	0.00
		(0.00)	(0.00)
N	72	72	74
R2	0.34	0.33	0.29

Appendix B

Industries' excess returns: Factor correlation

***ERS is a variable defined as *Excess Return Stock* and is only employed in R for the purpose of analysis.**

Factors correlation

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
CMA	1.000	-0.065	0.979	0.029	-0.022	-0.002	-0.988	0.957
ERM	-0.065	1.000	-0.071	0.145	0.104	-0.084	0.068	-0.036
HML	0.979	-0.071	1.000	0.025	-0.020	0.015	-0.942	0.902
RDP	0.029	0.145	0.025	1.000	-0.036	0.051	-0.051	0.065
RDRA	-0.022	0.104	-0.020	-0.036	1.000	0.039	0.022	-0.011
RDRG	-0.002	-0.084	0.015	0.051	0.039	1.000	0.004	-0.010
RMW	-0.988	0.068	-0.942	-0.051	0.022	0.004	1.000	-0.974
SMB	0.957	-0.036	0.902	0.065	-0.011	-0.010	-0.974	1.000

Automobiles and components (2510)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.007	0.447	0.001	-0.089	-0.024	-0.076	-0.001	-0.053

Banks (4010)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	-0.033	-0.051	-0.010	-0.183	-0.055	0.014	0.056	-0.266

Capital goods (2010)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.023	0.553	-0.009	-0.151	0.074	-0.201	-0.026	0.022

Commercial and professional services (2020)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.129	0.568	0.093	-0.120	0.000	-0.325	-0.132	0.117

Consumer Durables and Apparel (2520)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	-0.018	0.163	-0.013	-0.070	-0.033	0.003	0.031	-0.009

Consumer services (2530)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.183	0.504	0.151	-0.108	0.009	-0.295	-0.183	0.177

Diversified financials (4020)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	-0.015	0.035	-0.013	-0.181	-0.055	0.029	0.065	-0.169

Energy (1010)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.046	-0.012	-0.149	0.032	-0.022	0.022	-0.182	0.206

Food & staples retailing (3010)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	-0.052	0.288	-0.046	0.220	-0.134	-0.030	0.054	-0.042

Food, Beverage and Tobacco (3020)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.025	0.588	0.014	-0.103	-0.108	-0.185	-0.027	0.017

Health care equipment and services (3510)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	-0.012	0.461	-0.018	-0.103	0.012	-0.145	0.022	-0.015

Household & Personal products (3030)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.004	-0.039	-0.189	0.066	-0.018	0.016	-0.143	0.159

Insurance (4030)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.063	0.472	0.047	-0.135	0.080	-0.125	-0.062	0.061

Materials (1510)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.012	0.562	-0.007	0.005	0.000	-0.313	-0.018	-0.018

Media & Entertainment (5020)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.021	0.556	0.008	-0.117	-0.016	-0.097	-0.018	0.005

Pharmaceuticals, Biotechnology and Life Sciences (3520)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.911	0.064	0.896	-0.007	-0.020	-0.058	-0.899	0.872

Real Estate (6010)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.080	0.414	0.080	-0.240	-0.015	-0.539	-0.065	0.058

Retailing (2550)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.021	0.596	0.005	0.060	-0.109	0.003	-0.008	-0.009

Semiconductors and Semiconductors equipment (4530)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.999	-0.074	0.980	0.030	-0.024	0.020	-0.988	0.953

Software and services (4510)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	-0.023	0.192	-0.040	-0.184	-0.013	0.000	0.022	-0.010

Technology Hardware and equipment (4520)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.011	0.152	-0.009	-0.052	0.015	-0.986	-0.012	0.016

Telecommunication services (5010)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	-0.008	0.102	-0.010	0.136	-0.005	0.012	0.013	-0.008

Transportation (2030)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	-0.021	0.271	-0.035	-0.102	0.151	-0.039	0.015	-0.011

Utilities (5510)

	CMA	ERM	HML	RDP	RDRA	RDRG	RMW	SMB
ERS	0.036	0.462	-0.012	-0.159	0.073	-0.187	-0.053	0.052