The Specialization Effect in Private Equity

-An Empirical Analysis

Knud Kloster and Sigurd Moksnes Røthe

Supervisor: Carsten Bienz

Master Thesis in Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Bergen - Spring 2014

Abstract

The objective of this thesis is primarily to document the effects of specialization on the success of private equity(PE) funds through an empirical approach. We examine the relationship between a PE-fund's degree of specialization, measured by the Herfindahl-Hirschman Index, and the fund's success measured by the portion of portfolio companies that go public through an initial public offering(IPO) or a trade sale. We find that degree of industrial specialization has a statistically and economically significant positive effect on a fund's IPO share, especially for funds that have experienced at least one IPO exit. This provides evidence strengthening the specialization hypothesis. We find that one of the key drivers to achieve positive IPO share in the first place is fund size. We also find a non-linear relationship between geographical specialization and divestments through M&As. We find no significant differences of the specialization effect between Venture Capital and Buyout funds. Neither do we find that funds with higher IPO share are related to more risk through higher failure rate. This supports the belief that the degree of specialization can be used to control portfolio risk.

Keywords: *Private Equity, Specialization, Hirschman-Herfindahl Index*

Preface

This thesis was written as part of our MSc in Economics and Business Administration at the Norwegian School of Economics (NHH) and corresponds to one semester of full-time studies.

The choice of working within the field of private equity and our focus on fund specialization was mainly due to our fascination of the Norwegian private equity investor HitecVision, and their success in the oil and gas industry. Knowing that their strategy revolves around industry specific investments inspired us to examine the existing literature on the specialization topic. Our supervisor suggested some related articles on the theme, and because of contradictory conclusions in these studies we wanted to contribute to the understanding of fund specialization effects.

Working on this thesis has been a great learning experience and the process has been both challenging and rewarding. The challenges faced have without doubt increased our problem solving skills and our ability to overcome technical difficulties. Our decision of working together was an easy choice, as we have very similar interests within financial economics and had a similar idea of what we wanted to make out of this study. We believe that our collaboration during these past six months have strengthened the value of our thesis through mutual discussions and complementary knowledge. Working together has enabled us to dig deeper into the subject. We hope that others will find our work as a positive contribution to the field of private equity research, and that we are able to broaden the reader's insight of the significance of fund specialization.

We would like to thank our families for their continued support throughout our studies. Ultimately, we would like to express our sincere gratitude to our supervisor Carsten Bienz for his valuable inputs and help throughout the process. We would also like to thank him for his open door policy and for taking the time to discuss different considerations with us.

Knud Kloster — Sigurd Moksnes Røthe

Bergen - June 17, 2014

Contents

1 Introduction

Since the breakthrough of private equity as an asset class during the mid 1980s, the industry has been growing considerably during the past three decades. Here in Norway, the PE industry has become more popular and is now assigned with more interest from business newspapers as well as the general public. The importance of private equity investing seems to have grown substantially for most institutional investors and pension funds, and the total assets under management of PE funds have been growing with an exponential rate which hit the $3 trillion mark at the end of 2011 (Kenyon, 2012). In Europe, the overall fundraising more than doubled from €24.6 billion in 2012 to €53.6 billion in 2013 (EVCA, 2014). As shown by these industry facts, there is without doubt a massive amount of capital that is currently flowing into the PE sector. Because of the increasing importance of private equity as an asset class, the amount of research on the matter has also been taking off. However, despite the growth of academic research, there are many important questions that have not been fully answered. This is largely due to the lack of reliable return data available to researchers. The data problem arises because most PE firms are exempt from public disclosure requirements, and each market participant has few incentives to disclose the performance of their fund unless it has been among the top performers and achieved above average returns. The idea of performance persistence within the private equity industry has been documented in studies by Kaplan and Schoar (2005) and this further enhances the reluctance to provide reliable information regarding fund performance. One of the greatest concerns of previous research on private equity performance has been the sample collection procedure and the validity of gathered data. The fear is that the sample suffers from a positive selection bias arising from the fact that only the best performing firms benefit from presenting their returns.

In this thesis we propose an alternative approach to assess the effect of fund specialization within industry or geographic areas on the portion of successful portfolio company exits. In this way we make use of data that are of less sensitive character for fund managers than those of return data. We believe that this makes our analysis more reliable and robust with respect to the truthfulness of PE firm responses, thus reducing the potential selection bias. Another presumed strength of our procedure is that we build our analysis upon information within the Thomson Reuters VentureXpert database. This data is relatively accessible for anyone associated with either a business school or PE firm and only requires a product license. By using the subroutine macros that we have developed in Excel, our results can be easily reproduced and challenged. They can also be used as a basis for organizing information from VentureXpert for the purpose of other research than what is considered here. Because our method applies non-sensitive and easily accessible information for target group readers, we consider problems originating from self-reporting by funds as a minor problem. Our method also relies upon some well documented assumptions of portfolio company exit preferences found in other research. We trust on the empirical findings that there exists a hierarchy of preferred exit strategies within the private equity industry. In compliance with this pecking order of exits, we use three approximations of fund success: namely the IPO share, M&A share and Success as the combination of the first two. In order to account for potential heteroskedasticity and autocorrelation in the residuals in our regression analysis, we take use of the clustered standard error procedure. This makes our results robust to the possibility that residuals of an Ordinary Least Square (OLS) regression violates the homoskedastic assumption and constant variance, because observations are believed to be correlated across both the firm and time dimensions. However, using OLS when modelling fractional response variables has its limitations. This especially applies when there are large outliers in the data that could result in predictions that lies outside the defined limits. Because of this our analysis is expanded with non-linear models that ensures that this logical breach will not happen. Using Generalized Linear Model (GLM) for Fraction Responses (Papke and Wooldridge, 1996) as a basis for these non-linear models, we look at the average marginal effects of our significant regressors on the fractional response variable. Many research papers do not emphasize the potential cluster correlation among observations, and failing to adjust for this could lead to severe under-estimation of standard errors and consequently over-rejection of the null-hypothesis using standard hypothesis tests.

Using the VentureXpert database as a foundation we construct a global dataset from 1975 to 2003, containing 1,042 General Partners and 2,898 PE funds. With this amount of observations, we feel that the data is more than sufficient for the purpose of making statistical inference and to document the effect of fund specialization. To increase the degree of detail in our data, we have included several variables from other sources. More specifically we have expanded the data set with publicly available information regarding the growth of global gross domestic product (GDP), Price - Earnings ratio, the MSCI World Index and U.S. Treasury Bond and Bill rates. These financial figures are used to instrument the financial climate and the attractiveness of IPOs.

Throughout the thesis we aim to describe the dataset in addition to the main fund characteristics. Most of our focus will be to elaborate upon how these data characteristics comply with our predefined hypotheses and the effects of fund specialization. In the end we will compare the results of our study with the findings of other research and discuss the impact and economic significance of our results. As mentioned above, we explain fund success by three different measures. Our strongest definition counts the portion of a fund's portfolio companies that underwent an Initial Public Offering, IPO share. Our results indicate that the higher the degree of specialization of funds with respect to industry, the higher the share of IPOs is expected to be on average. The effect of Industry specialization also holds for the combined success measure. Our results are statistically significant on the 99% level and robust to several different regression methods.

We also find determinants of fund success similar of those found in studies considering the internal rate of return and other quantitative measures of fund performance. Kaplan and Schoar (2005) suggest a concave relation between fund size and performance. I.e. that larger funds have higher public market equivalents, but that performance declines as they become larger. Our results support this relation to some extent. Our analysis suggest that increasing fund sizes have a significant impact on the probability of achieving positive IPO shares, but we do not find that increasing the fund size has a significant impact, neither statistically nor economically, on IPO share for funds with at least one IPO exit. We also find that the investment opportunities approximated by the Price Earnings ratio of the S&P500 over the 10 year period from the fund vintage year is positively correlated with the percentage of IPOs. Also the 10 year US Treasury Bond Yield, the 10 year geometrical average return on the MSCI World Index is positively correlated with the share of IPOs and the share of successful exits. Thus we find that a large portion of the PE fund's success is determined by exogenous factors such as the development and situation in the public stock markets which are beyond the control of

General Partners (GP). Another result that we found is that funds of firms located in Europe and North America, on average, have a lower degree of IPO share than those of firms that are located in other parts of the world. This is perhaps the most surprising of our findings, and something that could be analysed in future studies. Our initial belief that the portion of IPOs in the western part of the world would be higher due to better quality, development and reputation of domestic financial institutions, is by the look of our data not correct. Doidge et al. (2011) show in their article *The U.S. Left Behind: The Rise of IPO Activity Around the World* that there has been a decline in IPO activity within the U.S. and that there has been a massive increase in other parts of the world. Another possible explanation can be due to the sample selection procedures at VenturеXpert. Our major concern regarding this result is that Europe and North America seem less attractive fund locations because the information gathered from the rest of the world consists of only the largest and best performing funds.

We test findings that have been shown by Gompers et al. (2005) that industry specialization of VCs tended to be more successful, on a global and non-investment stage specific scale. Our results are in compliance with the specialization hypothesis (Lossen, 2006). However our results contradicts the evidence provided by Ljungqvist and Richards (2003) and Brigl et al. (2008) who came to the conclusion that industry diversification or specialization has no significant effect on returns or the firm performance. In addition to documenting the effect of industry specialization on fund exit success, we examine the possibility of differences of this effect between VC and Buyout firms. According to our results we find no statistical evidence that the effect of specialization within industries differs between the two segments.

The remaining part of the thesis is structured as follows: Section 2 will provide a short presentation of this thesis' underlying theory and will mostly serve as an introduction to our final formulation of hypotheses. Section 3 describes the initial sample data, the data processing tasks and the variables that make up the dataset. The 4th section provides a detailed description of our most important econometric considerations and methods and section 5 continues with discussion of the results of the analysis. All figures and regression tables are located in the appendix.

1.1 Background

This section will provide an introduction of the private equity industry and the PE model. As a thorough discussion of the private equity components would be far too comprehensive, we limit this section to a brief description necessary in understanding the context of our thesis. For a more complete discussion of the private equity model and its components, please refer to "*Private Equity Demystified*" by Gilligan and Wright (2010).

Private equity includes investing capital in unquoted companies or public companies that are delisted as part of the transaction (Gilligan and Wright, 2010). The objective is to generate an optimal risk-adjusted rate of return on investments through an active ownership. The primary rewards of PE investments are capital gains realized by the exit of investments. Private equity investments are split into different segments according to the portfolio company's stage in the business life cycle. The literature and PE industry itself distinguish between several segments, such as Venture Capital, Mezzanine, Distressed and Buyouts. For practical reasons, one often apply a broader definition of PE segments, where it is only split into Venture capital (VC) and Buyouts (BO) (Müller, 2008).

Financing companies through PE investments serves two purposes. First, private equity is used to finance growth by providing capital in order to develop new products, expand operations or make acquisitions. This is typical for venture capital investments. Secondly, private equity can resolve ownership and management issues by providing professional expertise through active participation at the company board level, or similar. Illiquidity is another aspect of PE investments. The portfolio companies are non-listed and privately held for a relatively long investment period (normally 10 years). In addition, there does not exist any active second hand market where partnership stakes can be instantly bought or sold. The sale of any limited partnership also requires that the General Partner (GP) approves the alternate investor. Investors have thus limited control over how their committed capital is invested during the fund's life time and their committed capital is therefore considered illiquid (Müller, 2008).

The typical private equity model is illustrated in Figure 1 in the appendix. The general partner manages n number of funds, where each fund is associated with m limited partners. The limited partners provide the majority of equity to the fund and are thus entitled to most of the capital gains. The total committed capital from investors and the amount committed of the GP, are in turn invested with a fraction of leverage in k portfolio companies.

Venture Capital - VC

Venture capital refers to investments in start-up and other early-stage companies characterized by strong business ideas with high growth potential and investment needs but low current cash flows. Due to the high risk of investing in early stage firms, they have limited financing alternatives other than equity financing provided by venture capitalists.

Buyouts - BO

Buyouts or late stage investments, consist of investments in more mature and established companies where the cash flows normally are positive and relatively predictable. This allows additional leverage in investments through debt financing. The potential returns to the equity providers of buyouts mainly originate from financial restructuring, cost reduction and efficiency improvements (Müller, 2008). Active ownership is the main key to success in buyouts, and it requires GP involvement and a strong collaboration between the portfolio company and the fund managers. The buyout segment usually involves acquiring the majority of shares in a mature company, and are often larger in size than venture capital investments.

Private Equity Fund

Private equity funds function as pooled investments by a number of investors in equity and equity-related securities of companies. These funds are typically structured as limited partnerships with a finite life time, normally ten years. These funds are called closed-end funds (Gilligan and Wright, 2010). An alternative fund structure is to establish a company. Whatever structure is chosen, the investors in PE funds commit capital to the fund and fund managers invest further in portfolio companies. When realisations occur, the fund repays capital to its investors and earnings are not usually retained.

General Partners - GP

The general partner in a limited partnership is responsible for all management decisions of the partnership. The general partner invests the fund's committed capital in public and private companies, manages the portfolio of investments and seeks to exit the investments in the future for sizable returns. A general partner may manage one or a few funds that may have different investment restrictions such as geography, industry or typical size of each investment. GPs typically invest a minority share of equity to increase the level of confidence in their funds, and typically charge an annual fee of around 1-2% of capital under management in addition to a share of approximately 20% of the fund's profits, called carried interest. The GPs have fiduciary responsibility to act for the benefit of the limited partners and is liable for all its actions.

Limited Partners - LP

Investors that commit capital to a limited partnership are called limited partners. Limited partners have limited liability and are not involved with the day-to-day operations of a PE-fund. The LPs commit a substantial portion of the fund's equity, and are thus entitled to most of the income, capital gains, and tax benefits. LPs generally consist of pension funds, institutional accounts and high net worth individuals.

2 Theory and Hypotheses

2.1 Related Research

In the past few decades the amount of research revolving around the private equity industry has increased substantially, but still a lot of questions are not clearly answered. The reluctance shown by PE firms to provide information leads to contradicting conclusions on what drives performance of PE funds.

The research that relates most to this paper is the work done by Lossen (2006), who according to himself was the first to perform a systematic analysis of the impact of diversification on the performance of private equity funds. His results show that there is in fact a negative relationship between industry specialization and performance, contradicting his own specialization theory. Still, there are several differences in our thesis compared to Lossen's. Lossen only use a dataset containing of 100 PE funds of 34 PE firms where 60% of the funds were still active, which is a rather small sample compared to ours. Furthermore, the sample collected was not random, and limited to funds that belonged to a PE firm that undertook fundraising between 2000 and 2005. In addition, Lossen relied on manually collected information regarding cash flows to measure performance. The validity of these figures could be questioned.

Gompers et al. (2005) provided more evidence on industry specialization by finding evidence that industry specialized VCs tended to be more successful[1]. Contrary evidence has been provided by Ljungqvist and Richards (2003) and Brigl et al. (2008) who concluded that industry diversification does not have significant effect on returns or the firm performance. Recent studies have examined possible relationships between the degree of fund-level specialization and performance, however results are inconclusive. Other research related to drivers of PE performance suggests that availability of debt financing is a major factor in determining private equity activity. Ljungqvist et al. (2007) report accelerating investment pace of private equity funds as interest rates decline. Kaplan and Schoar (2005) and Gottschalg et al. (2004) report positive and concave relationship be-

[1] The authors define specialization as the fraction of follow-on funds that invest in the same industry as the first fund. This is in our view a bit misleading or at least uninformative as it does not elaborate on the varieties of those industries that differ from the first fund.

tween fund size and the rate of return of PE funds. The concave shape of the relationship suggests that there is a positive but decreasing marginal effect of increasing fund size. Their results also suggests a positive impact of firm experience on the rate of return of PE funds.

Diller and Kaserer (2005) and Gottschalg et al. (2004) report a negative influence on a funds rate of return from the annual rate of return of the MSCI Europe Index. They also find that an increasing amount of new funds raised by the PE industry in the vintage year of a fund, decreases fund returns.

2.2 Theory

As mentioned in previous sections, the amount of research has grown immensely. Private equity literature now covers a wide area of topics regarding PE returns, the role of corporate governance and the operational effects of PE investments. However, because private equity is largely exempt from public disclosure requirements, the PE data gathered in research studies have been somewhat unreliable due to survivorship bias and reporting bias, leading to contradictory conclusions between papers.

The main objective of this thesis is to provide empirical evidence to assess whether the degree of fund specialization has a significant impact on their success. The thesis will also contribute in determining other important determinants of fund performance. Although we would prefer to measure fund performance directly through return data, we acknowledge the industry's general reluctance to provide such information, and have therefore chosen an alternative approach by measuring fund performance indirectly. Our approximation of fund performance relies on the type of exit chosen for a portfolio company, and is divided into three categories ranked by the exit preference. Applying exits as indicator of fund success and performance is largely supported by other research. Bienz and Leite (2008) provide empirical evidence that more profitable companies will go public while less profitable companies with a higher need for oversight and monitoring will be sold through a trade sale[2]. Gompers (1995) refers to Venture Economics' review of returns on venture capital investments, and concludes that VC funds that eventually perform an IPO yielded the highest return for venture investors. We define successful exit strategies in two degrees, namely IPO and Mergers & Acquisitions, where IPO exits are ranked as the better outcome and will therefore be prioritized in this thesis. We later combine these two to analyse overall successful exits (IPO and M&A). We believe that using data of less sensitive character to PE funds, benefits the robustness of our dataset and reduces the scale of potential selection bias. In the following subsections we provide the theoretical background leading up to our three hypotheses.

2.2.1 Specialization Hypothesis

Research made by Bottazzi et al. (2004) and Gompers et al. (2005), identified two main strategic dimensions of PE firms. Namely, firm investment focus by industry and by stage. Cressy et al. (2007) argue that PE firms specialized relative to their competitors, possess a deeper knowledge regarding the competitive environment of their portfolio companies, and therefore have an advantage when assessing the strengths and weaknesses of potential portfolio companies. On this basis, specialized PE firms are believed to have an advantage in selecting superior investments by picking better performers with larger potential, as well as providing more effective monitoring and advice than others.

According to Lossen (2006) there are at least three advantages of being specialized in particular financing stages, industries and countries. First, before investment decisions are made, PE firms must go through a multi-staged selection process, in which the PE firm faces substantial information asymmetry as opposed to the management team or current owner of the portfolio company. According to the specialization hypothesis, specific knowledge in technology and business of the potential investment will reduce the information gap and thus improve the PE firm's ability to assess the attractiveness of the potential investments.

[2]We acknowledge that this not necessarily equals that IPOs are a more profitable divestment strategy, but it does indicate that if a company goes public, it is highly profitable and it is reasonable to assume that the PE-fund would only perform an IPO of a portfolio company if it gave a positive return on the investment

The second aspect relates to the financing of portfolio companies. During financing, a principal agent problem arises between the PE firm (principal) and the portfolio company (agent). There might be a difference of interest between parties and undesirable pursuits of individual goals might occur on the agent's behalf. This will in turn raise agency costs for the GP. Ewens et al. (2013) argue that, due to principal-agent problems, GPs necessarily hold undiversified positions. Thus, part of the compensation to PE firms relates to the level of idiosyncratic risk faced by fund managers.

Thirdly, PE firms claim to be able to add value to their investments by involving themselves in managerial activities in portfolio companies. If the PE firm has a lot of know-how about the portfolio company's activities the more value adding services they are likely to contribute, making the investment more profitable in theory. Empirical studies on the economic effects of buyouts transactions find that PE firms add value to their investees by the provision of long-term investment capital, better corporate governance and quality advice, thus improving their operational efficiency (Jensen, 1989; Kaplan, 1989).

In contrast to traditional finance theory which suggests that portfolios should be diversified over several companies and different industries to minimize unsystematic risk, private equity funds seem to benefit from the opposite. The degree of specialization in a fund can be used to control portfolio risk (Norton and Tenenbaum, 1993), and in addition to reducing risk they also benefit from exploiting relevant knowledge and business insight in their area of specialization. Thus making higher returns. The specialization hypothesis proposes a negative relationship between the level of diversification and the rate of return of PE funds. Accordingly, it suggests a positive relation between the level of diversification and the percentage of loss in a PE fund (Lossen, 2006).

2.3 Hypotheses

Our primary objective of this thesis is to examine the determinants of successful portfolio company exits and mainly determine the PE firm specialization advantages. In related studies considering the specialization hypothesis, it has been suggested that specialization across stages, industries and geographical areas have positive influence on fund success and thus also returns. As mentioned in the section 2.1, the evidence of specialization advantages has been inconclusive. Because of this inconsistency we aim to provide further empirical evidence of the specialization hypothesis. We initially wanted to examine all three areas of specialization, but due to data limitations in VentureXpert, we were left unable to discover stage information of individual portfolio companies. In addition to examining the specialization effect on fund success, we also study potential differences of the specialization effect between VC and BO funds. As a final area of interest, we have studied the determinants of fund failure rate and compared significant variables of failure to those of our primary measure of success, IPO. On the basis of this comparison of coefficients we wanted to discuss whether variables related to successful exits also would be related to higher risk, instrumented by the share of failed investments within funds.

2.3.1 Hypothesis I: Specialization Effects on PE-fund Success

As our main attempt of this thesis is to study the effects of fund specialization on fund success, and due to the fact that we have not gained access to fund return data, we have defined success in three different levels: IPO share, M&A share and Success. The last measure is the combination of the two previous success measures. We noticed that studies using return data often had small sample sizes and remarked that their return data could be subject to positive biases.

Based on the arguments of the specialization hypothesis, we want to see if there is empirical evidence that funds specialized on either industries or geographical areas are able to drive higher operational performance and achieve larger percentage of successful exits. Does deeper business knowledge and expertise on specific industry areas really drive fund exits to be more successful, or does the increased number of investment opportunities of a generalist fund outweigh this element?

$H1$: How does the degree of specialization affect the success of PE-funds?

2.3.2 Hypothesis II: Differences in Specialization Effects on IPOshare between Venture Capital and Buyout Funds

While researching the specialization topic, we found that the samples in several other articles often were limited, either in terms of investment stage, time period, or geographical areas. Because our sample consists of both VC, BO and Other PE firms, we wanted to examine whether there is a difference between the effect of specialization on fund success between these segments. In the majority of related research, the samples have been limited to examine specific investment stages or geographical areas and very few have studied the effect of specialization at an overall and global scale. Thus we find it interesting to take use of the width of our data set to examine the potential differences in the advantages of fund specialization across fund investment stages.

$H2$: Does the specialization effect differ between Venture Capital and Buyout?

2.3.3 Hypothesis III: The Degree of Specialization and Fund Failure Rate

The final aspect we wanted to take a look at was the number of failed deals within funds. We already had the amount of failed deals in our data and we therefore found it interesting to use this information in order to consider what drives fund failure. The failure rate also functions as a proxy for risk, and it would be interesting to see whether determinants of fund success (and implicitly returns) also are related to higher risks.

$H3$: Does the degree of fund specialization relate to increased risk, measured by fund failure rate?

3 Sample and Data

The data in our analysis is retrieved from a few separate sources including Thomson Reuters VentureXpert, Datastream and US Federal Reserve. The primary source of information is VentureXpert, which contributes private equity firm and fund information. The unaltered raw data consist of 1,241 General Partners in all investment stages, managing 5,497 PE-funds, with a total of 60,784 unique fund and portfolio company combinations. However the initial information from VentureXpert is not organized, and returns are only available on an industry aggregated basis. To make the dataset compatible with statistical software, such as Stata, we needed to do a preliminary structuring process as explained in detail in section 3.1. In addition to the raw and the structured dataset, a Visual Basic for Applications (VBA) macro is attached in the appendix of the electronic version for the convenience of replicating our dataset. Another problem of the VentureXpert database is that there are missing information for certain funds' variables. After the adjustment, our sample is reduced considerably. The organized dataset contains information of 1159 General Partners and 3,902 PE-funds. A final concern is the degree of quality on observations made before 1975. Gompers and Lerner (2004) state that the funds whose documents are in the VentureXpert files appear to represent a random sample of the industry. But they also argue that since information gathering first started on a regular basis in the late 1970s,"*...the occasional information prior to this do not appear to be gathered systematically*". We acknowledge these concerns and thus restrict our analysis to funds with a vintage year from 1975 to 2003. The upper limit restriction is made due to the investment cycle of PE-funds. Recent funds have not yet operated long enough for measuring their lifetime performance. In this way we believe that our dataset becomes more robust and functional.

The additional information necessary for our analysis was retrieved from what we consider as reliable sources. The MSCI World Index and annual World GDP growth was retrieved from Thomson Reuters Datastream, the market yield on U.S. Treasury securities from the Federal Reserves database (FED, 2014), and historical Price/Earnings ratios were gathered from *Irrational Exuberance* by Schiller (2000).

As opposed to many other empirical studies, our data is not restricted to observations from only one investment stage, e.g. Venture Capital, or to a specific region, e.g. Europe or the U.S. We believe that this strengthens the relevance of our thesis by allowing for more observations and the possible findings of differences across geographical regions and investment segments.

Survivorship Bias - SB

The data is reported based on realised transactions that pay out cash to limited partners. In situations of survivorship bias, the reported financial figures and the recorded sample of the population would be positively biased and consist mainly of companies that perform better and that remains in the market over the observation period. Worse performers that carry out write-downs or write-offs tends to delay such actions towards the time of liquidation of a fund rather than doing this on a regular basis. This leads to portfolio investments that should be written off, are kept in the portfolio as "living deads" and not recorded in transaction databases. This means for instance that the average holding periods could be underestimated and include less successful investments which are not revealed. Even though we expect that there could be a survivorship bias present, we do not consider it to be severe because of the less sensitive nature of the data used in our analysis. In addition to this, there is few options available to adjust for this potential survivorship bias and we must assume that it does not affect the analysis.

3.1 Transformation and Data Processing

A problem with VentureXpert is that the output report is an Excel spreadsheet where details regarding the funds and their respective portfolio companies are lumped together in single cells. This makes it impossible to perform operations in Stata as the software is not able to separate this information automatically. Another problem that occurred was two different funds that went by the same name, but was written differently in terms of capital letters, more specifically Bancboston Ventures vs BancBoston Ventures. This produced a lot of errors as VBA would treat this as two different funds while Excel did not. To solve these and several other issues we wrote subroutine macros in Visual Basic for Applications (VBA). The macros serves multiple purposes. First and foremost the macros are necessary in order to split and organize data. Secondly, writing macros automated the calculation procedures and enabled us to work with large amount of data.

The automation is also valuable for replication and peer review purposes. Thus we have provided all the programming codes in the appendix of the electronic version of this thesis [3].

The bottom line of the macros' function is to split and sort all the available information into separate cells, so that the data structure is in compliance with Stata or any other statistical software, as well as aligning them with their respective funds and managing firms so that all rows are correct and complete with no missing values. We removed all funds with no numeric values, as these would not provide us with any useful information in our studies and would have been omitted anyways when running regressions. The macros also consolidate data retrieved from different sources, namely figures based on MSCI world index, global GDP and US interest rates with the metadata from VentureXpert.

We have also written macros to compute all the numeric figures that we found useful or interesting to test in our regressions. This also included the Herfindahl-Hirschman Index (HHI) which measures industrial and geographical concentration of a fund, based on its investments in portfolio companies. The automation process makes it a lot more time-efficient, reproducible and less prone to human calculation errors. The programming and debugging process has been tedious work, but the end result is more robust and it has made it possible to easily adjust for more parameters or calculating methods if needed. Manipulation of the large sample would simply not have been possible in the time limit given if we were to do it manually in excel. After all alterations had been done, we then sort and copy all the data we want to investigate further in to a new single worksheet that is finally imported into Stata. Because of this preliminary work there will be no need for additional sorting and merging operations of the dataset in the statistical software.

[3] The macros are demanding to run due to the large number of observations and simultaneous operations across separate spreadsheets, so a relatively powerful computer (in 2014) is required to prevent Excel from "crashing". Specifications of the computer used for data manipulation include 8 GB ram and an Intel i7 quad-core CPU @ 2.20GHz

3.2 Explanation of Important Variables

3.2.1 Specialization

To measure specialization we decided to use the Hirschman-Herfindahl Index (HHI). A commonly accepted measure of concentration, used by anti-trust agencies, among others, to assess the degree of market power. We calculate the HHI using the following formula:

$$HHI = \sum_{i=1}^{n} (\frac{s_i}{S})^2$$

Where s_i is the amount invested by the PE fund in industry or country i, and S is the total amount invested by the fund in all portfolio industries or countries combined.

The HHI-score is a value ranging from 0 (infinite diversification) to 1 (perfectly concentrated). As some funds are only listed with one portfolio company we would expect to find a high number of observations with a value of HHI = 1.
To calculate the industrial HHI, $HHIIndustry$, we have categorized the industry of the portfolio company based on 19 unique categories. This categorization is not too general nor too specific in our opinion, as it separates the industry into 19 different areas of expertise, while still holding similar industries in one bracket. The different industries are Construction, Manufacturing, Consumer Related, Industrial/Energy, Computer Software, Communications, Medical/Health, Internet Specific, Business Services, Semiconductor/Electronics, Transportation, Biotechnology, Computer Hardware, Financial Services, Agriculture/Forestry/Fish, Computer other, Utilities and Other. The HHI figures are calculated on the basis of 60,784 investments made by PE funds.

In addition to the industry concentration of a fund, we also calculate a geographical concentration index, $HHICountry$, based on the country in which a portfolio company is registered.

3.2.2 Success

To measure success we will use two different exit strategies as our dependent variables, namely $IPOshare$ and $MAshare$ as well as the sum of these two which results in our

Success-variable. *IPOshare* is the percentage of a fund's portfolio companies that had an Initial Public Offering (IPO) and is considered in this thesis to be the most successful outcome of exit strategies. *MAshare* on the other hand, is the percentage of a fund's portfolio companies that was either merged with or acquired by another company. We still view this exit strategy as successful, but less so than IPO exits. Lastly we will investigate the overall *Success*, by summing the share of portfolio companies that was exited through both IPOs and M&As.

3.2.3 Macro-Economic Variables

Literature suggests that capital flows into private equity as well as other assets when good market conditions are present. When demand increases, so does the prices and valuation which in turn would increase the incentives to go public due to higher pricing. To account for this *money chasing deals* phenomenon we include several macro-economic variables. Some of these variables depict macro-economic events during the life cycle of a fund, which usually last around 10 year. The variables we have chosen to include is growth in global Gross Domestic Product (*GDP10yrGeoAvg*) as a proxy for world production, as well as different instruments for financial factors such as the change in the MSCI World Index(*MSCITenYearFwdAvg*), the interest rates on 10 year US treasury bonds (*Bondyieldtenyear*), and Schiller's inflation adjusted Price Earnings ratio (*PE10*) for the S&P500. All these variables, except for Bondyieldtenyear which already reflects the future expectations, are measured by taking the geometric average for a 10 year period starting at a fund's vintage and going forward.

3.2.4 Omitted Variable Bias - OVB

Multiple regression is powerful in examining the effects of variables on which we have data. A major concern in corporate finance studies is that desirable data often are unavailable or impossible to quantify. Inability to observe these determinants means that instead of being amongst the explanatory variables, X, the omitted variables are included in the error term u. If data is unavailable, we cannot account for the omitted variables in the regression and the OLS estimators of the regression coefficients could be misleading. In other words, the mean of the sampling distribution of the OLS estimator in case of omitted variable bias might not equal the true effect on the dependent variable of a unit change in the regressor. If omitted variables are uncorrelated with the included regressors there is no problem, but correlation between the error term and the independent variables prevents us from making any causal interpretation (Roberts and Whited, 2012).

One possible solution to account for unobservable data is to use panel data regression. Using this procedure, one can control for some kinds of omitted variables without actually observing them. The idea behind the approach is that by studying the changes in the dependent variable over time, it is possible to eliminate the effect of omitted variables that differ across entities but are constant over time. However, this approach strictly requires that the sample in fact is panel data. In our case, we cannot uniquely identify observations across firm and vintage year because some firms have several funds within the same year. This results in several duplicating identifiers. In addition to this, some firms only have one fund eliminating the possibility of observing the effects over time and panel data requires that the entire population is recorded at least twice. Even though our sample has some similarities to panels, it is defined as pooled cross-sectional data. Because of this dissimilarity, using panel data regressions is not a possible solution to mitigate the potential omitted variable bias in our case.

In situations where an important control variable is missing from a regression, Judge et al. (1985) showed that omitted variable bias may lead to inflated coefficients if independent variables are correlated with the omitted variable. This effect may well be happening in our data, and we are not able to detect whether this leads to a bias or not. Because of these statistical problems we have considered potential instruments for unobservable factors we believe is relevant in explaining fund success. Instrumental variables should be correlated with the success of PE funds, but otherwise unrelated to the other regressors. During our assessment of this subject, we have not come up with any obvious variables that we have left out of our analysis. However, we have considered that it might have been useful to include whether the PE fund is independent or affiliated with e.g. bank, government or other institutions. The reason for our interest in this information is that this could detect whether the suggestions in the literature that non-independent

PE firms have fewer incentives to perform, and that this would lead to lower rates of success. One of the studies supporting this belief is Bottazzi et al. (2004), who conclude that independent VC firms are more active owners than others and that their portfolio companies are more likely to perform. Another potential variable of interest would be to include the possibility of CEO changes. A dummy for CEO change could probably function as an instrument for poor performing funds. Poor performance is most probably related to incapable CEOs or poor management in general, therefore changing the CEO may lead to an improvement in company performance.

To summarize, one can never be completely certain that regressions are free from OVB, and neither can we. We have attempted to include all variables we find relevant, and those believed to have an impact on *IPOshare*, *MAshare* and *Success*.

4 Methodology

This section will discuss the methodology and theoretical rationale for the different statistical considerations used in parts of the upcoming analysis. However we will not explain in detail the procedure for obtaining the OLS estimates, as our thesis emphasize on the application rather than the computation of regression coefficients, but all the steps of the regressions can be located in our Stata do-file.

4.1 Clustering

Wooldridge (2012) explains clustering as computing standard errors and test statistics that are robust to any form of serial correlation (and heteroskedasticity). Using clustered standard errors allows heteroskedasticity and autocorrelation in a way that is consistent with the assumption that $(X_{i1}, X_{i2}, ..., X_{iT}, u_{i1}, u_{i2}, ..., u_{iT}) \qquad i = 1, .., n$ are i.i.d. draws from a joint distribution. There are two common forms of dependence that are often present in finance applications. The residuals of a given firm may be correlated across years for a given firm (time-series dependence) or the residuals of a given year might be correlated across different firms (cross-sectional dependence) (Petersen, 2008). Omitted factors that enter the error term could have common elements for a firm's funds. Thus it is not reasonable to assume that the errors of funds are independent within the same firm, even though they are independent across firms. In other words firms are natural clusters, or groupings, of observations where u_{it} is correlated within the cluster, but not necessarily across clusters. The methods of clustering can be done in one dimension, e.g. firms, or in two dimensions to allow for clusters across both firm and time. It is even possible to account for multiple dimensions as proposed in the article Multi-Way Clustering by Cameron et al. (2006). Clustered standard errors are valid even when there is no heteroskedasticity, autocorrelation, or both in the data structure (Stock and Watson, 2011).

Adjusting for clustering is important, as a failure to do so can lead to under-estimating the standard errors of the coefficient estimates. Small standard errors lead to large t-statistics, and one might see statistical significance even when it does not exist. This could lead to consequent over-rejection of the H_0 using standard hypothesis tests (Cameron et al., 2006). Since clustered (non-independent) errors may produce incorrect estimates it can prevent us from conducting accurate statistical inference in our empirical research. The literature propose a vast amount of methods for adjusting standard errors to within-cluster correlation, and the appropriate method depends on the data (firm and/or time effect) and the intended use of the results. Because of the large amount of possible procedures it is good practice to report standard errors estimated by multiple methods. Comparing results from different methods can also help us diagnose potential problems with the model (Petersen, 2008).

In general we have that OLS standard errors are consistent if the regression residuals are uncorrelated across both firm and time, but this is often unlikely in practice. For example a market shock will cause correlation between firms at a moment in time, and persistent firm specific shocks will induce correlation across time. Other persistent common shocks, such as business cycles, can result in correlation between firms in different years (Thompson, 2010). A specific concern in our analysis is that the amount of IPOs are cyclical, and the IPO activity increases at peaks in the market for equities.

One dimensional methods allow us to neglect the form of the correlation within the cluster. The cost of this is that the residuals must be uncorrelated across clusters. This makes one dimensional methods less robust in situations where we have reason to believe that there are both time and firm effects present. A solution to this issue used by many empirical finance researchers is that one can parametrically estimate one of the dimensions by including dummy variables, e.g. dummies for each time period to absorb the time effect and then cluster by firm. Including time dummies completely remove the correlation between observations in the same time period, and we are then left with only a firm effect. This approach only works when the dependence is correctly specified. If the time effect is not fixed, then time dummies will not be able to remove the dependence completely and even standard errors clustered by firm can be biased (Petersen, 2008). Another limitation of the parametric procedure is that they limit the kind of covariates that can be included in the regression. If time dummies are used, then we cannot include macroeconomic variables, since they are collinear with the dummies (Thompson, 2010). The more sensible thing to do, when there is uncertainty about the precise form of the dependence, is to cluster on two dimensions simultaneously, e.g firm and time, using Mitchell Cameron's dual clustering method [4].

Thompson (2010) argues that double-clustering is most useful when the number of observations in each dimension is not too far apart. His point is that, all else equal, it is more important to cluster along the dimension with fewer observations. He also gives a stronger statement: *"if the dimensions are extremely unbalanced, we do not need to double-cluster at all, and only consider the parametric approach."*. We do not necessarily agree with the latter, because Mitchell Petersen (2008) argues that when there are a sufficient number of clusters in each dimension, standard errors clustered on multiple dimensions are unbiased and produce correctly sized confidence intervals whether the firm effect is permanent or temporary. Due to different consensus among researchers, we will proceed with both dual clustering and the parametric approach as a basis for our models.

4.2 Dual Clustering Procedure

In the following we explain the cluster2 program by Mitchell Petersen, which we apply to adjust for the possible autocorrelation and heteroskedasticity within both the firm and time clusters of our data. The source code can be obtained from Mitchell Petersen's own website and is also included in our Stata do-file. As a thorough mathematical explication of the procedure is beyond the scope of this thesis, we advice readers who desire a more formal presentation to Cameron et al. (2006). We will on the other hand explain the basics.

In the two-way clustering case, the program starts by doing a one-way clustering regression with respect to the first cluster variable, which in our case is the firm dimension. This gives us the White standard errors which are robust to within-cluster correlation. Clustering on firm, allows observations of the same firm to be correlated across years, but this assumes that observations between firms are uncorrelated in each year. The same kind of regression is then performed for the second cluster variable, specifically the year dimension. After each regression, the variance-covariance matrix with respect to each

[4]Which can be used in stata by implementing the cluster2 function. The .ado-file can be retrieved at http://www.kellogg.northwestern.edu/faculty/petersen/htm/papers/se/se_programming.htm

cluster is stored for the purpose of later calculations of the two-way clustered standard errors.

The third step of the program first checks whether the observations are unique in both cluster dimensions, and then it returns a value of 1 if there is only one observation per firmcluster-timecluster. It then makes a sum of the amount of unique observations and assess whether the average is 1 or not. If the average equals 1, then all observations must be unique. Satisfying this condition, the program runs a regression where the two-way cluster is simply the sum of the firm dimension and time dimension indicator matrices. If there are multiple observations in the firmcluster-timecluster, the computed average will not equal 1, and it then subtracts the $S^{G \cap H}$ matrix, to ensure that observations are not counted twice.

The essence of the dual clustering method, relies upon this adjustment.

$$S^{GH} = S^G + S^H - S^{G \cap H}$$

Where S^G, S^H and $S^{G \cap H}$ are the three different cluster-robust variance matrices for the estimator by one-way clustering in, respectively, the first dimension, the second dimension, and by the intersection of the first and second dimensions. The intersection variance matrix, $S^{G \cap H}$, is an $N \times N$ indicator matrix with ij^{th} entry equal to one if the i^{th} and j^{th} observation share any cluster, and equal to zero otherwise. This occurs when one firm has multiple observations, i.e. multiple funds with the same vintage year.

4.3 Alternative Methods to OLS

Because our proportion response variables only exist in the interval [0, 1], we should take the bounded nature into account when handling the data. The main problem of a linear model is that it might generate predictions outside the unit interval. This especially happens in situations with extreme values of the regressors, where the predictions are likely to be nonsensical and exceed the limits of the response. Another problem with OLS when modelling fractional responses, is that the estimated coefficients are constant in the entire interval, which is not a reasonable assumption for dependent variables within the unit range. More often in finance there is a nonlinear relationship where there are changing marginal effects throughout the range. The statistics literature proposes several different approaches for modelling proportions, but there is little consensus on which models are appropriate in different situations.

Despite that the application of OLS on fraction responses might be inappropriate, the most common practice of researchers is to apply OLS to their data (Kieschnick and McCullough, 2003). This is because OLS is by far the easiest model to estimate and interpret, or due to the authors' lack of knowledge about better alternatives. In situations where most of the observations are found in the middle, OLS might give good estimates. But when there are more heavy tails or mass points at the boundaries, the OLS becomes less accurate and attractive. A traditional solution to modelling fractional responses is to perform a logit transformation on the data. The transformed response variable, y*, is obtained through this relationship (Baum 2008):

$$logit : y^* = ln(\frac{y}{(1-y)}) = X\beta + \epsilon$$

By considering the above equation, one immediately realize that using logit is only appropriate when the dependent variable is strictly between 0 and 1. Observations at the unit interval limits fail the transformation and would subsequently be dropped from the estimation sample. Based on this we do not consider logit to be appropriate for our data.

A more suitable method is the Generalized Linear Model (GLM) for fractional responses as proposed by Papke and Wooldridge (1996). This method handles the bounded nature of the data and take into account both the zeros and ones as well as the intermediate values. The approach specifies a quasi-likelihood regression model for continuously measured proportions with a finite number of boundary observations. The approach generalizes linear regression by allowing the linear model to be related to the response variable via a link function, which often is set to be the logit transformation. It also take use of a distributional family where one make assumptions about the distribution of the dependent variable. Baum (2008) states that the binomial distribution is a good choice of family even if the response variable is continuous. The fractional regression approach stipulates that the variance of the binomial distribution must go to zero as the mean goes to either zero or one, because the variable approaches a constant in these two situations (Baum, 2008). The drawback of the fractional GLM approach is that it does not allow for an alternative model to generate the observations found at the limit points. If different factors generate the observations at the limits, a sample selection issues arises. In our situation of IPO share, we have a mass point at zero with 1,014 observations, and there might be a different process generating the extreme values for our success responses, than what underlies the continuous part.

When modelling a fractional dependent variable it is often a good idea to look at a frequency table. A marked spike at zero and/or one may raise doubt about a single model fitted to all data. On this basis, we considered a two-part fractional regression model as suggested by Ramalho and Silva (2009) and Ramalho et al. (2011). The distribution of our dependent variables can be found in Figure 2 in the appendix.

Considering the graph of IPOshare in Figure 2, it seems that observations at the zero boundary occur with too large frequency than what seems to be efficient with a simple model. A better approach will be to employ the two-part alternative. The discrete component of the approach is modelled as a binary model, and the continuous component as a fractional regression model as suggested by Papke and Wooldridge (1996).

The first part of the Two-part fractional regression model, hereby referred to as the two-part model, governs participation: i.e. specifies a binary outcome model to explain the probability of a firm having successful exits or not. The model define

$$Y^* = \begin{cases} 0 & Y = 0 \\ 1 & Y \in (0,1] \end{cases}$$

And then,

$$Pr(Y^* = 1|X) = Pr(Y \in (0,1]|X) = F(X\theta)$$

Where θ is a vector of variable coefficients and $F(\cdot)$ is e.g. the cumulative logistic distribution function. The logit model is then estimated by maximum likelihood using the whole sample. The second part of the two-part model handles the positive outcomes, in other words all funds with IPOshares that are not equal to zero. In the second part, the procedure applies a function $G(\cdot)$ similar to the one defined earlier.

$$E(Y|X, Y \in (0,1]) = G(X\gamma)$$

This second, continuous part, can then be estimated by quasi likelihood estimation. For more detailed information regarding this procedure, please refer to Ramalho and Silva (2009).

Kieschnick and McCullough (2003) recommend that researchers use either a parametric regression model based upon e.g. a beta distribution or a quasi-likelihood model as proposed by Papke and Wooldridge (1996). In small samples they prefer parametric approaches and quasi-likelihood approaches are appropriate in large samples, so that the asymptotics of the likelihood estimation can be justified. A weakness of the GLM model with logit link and other non-linear models is that the coefficients are not as easily interpreted as those of a linear regression. Since these models are non-linear, the marginal effects of the coefficients will also be non-constant. Consequently, if the value of one independent variable is changed or an additional variable is included, the marginal effect will change as well. When presenting the marginal effects on the dependent variable, one can mainly choose between two methods: the partial effect on the average (PEA) or the average partial effect (APE) (Woolridge, 2009). Because we have dummy variables in our regressions, the use of PEA becomes meaningless. The reason behind this is that the average of the dummy could be a decimal, but the only possible values one can observe for the dummy, is binary. Thus the average value of the dummy cannot correspond to the value observed for the average fund. Because of this we decide to use the APE when presenting our results in the upcoming analysis. In this method a coefficient represents the average marginal effect for all the values of the corresponding explanatory variable in the sample.

5 Results

This section provides a discussion of the results from the analysis made on the dataset of global PE funds. The first part will discuss the regression results related to our first hypothesis regarding the determinants of fund success and especially the effect of fund specialization. This part is subsequently divided into three sections in accordance with our three measures of fund success. Each individual part will elaborate on the economic effects of our findings. The second part includes a discussion of results with respect to the second hypothesis. In this part we will discuss whether there is empirical evidence in our data that there is a difference in the direction and magnitude of industry specialization between venture capital and buyout funds. Our third, and final section, will discuss the failure rate of funds and aims to reveal whether more specialized funds are related to a higher degree of failure. This section's objective is to see whether the determinants of being more profitable, i.e. higher shares of IPOs, also have a larger portion of failed exits which indicates that higher returns are related to higher risk.

5.1 Pre-Regression Findings

Before running regressions on the data set, a correlation analysis was performed in order to detect linear relationships and to assess possible related problems. Variables with high correlation coefficients express a potential violation of the OLS assumption of no perfect collinearity. The analysis was also used as basis for excluding variables that are accounted for through high correlation with other measures. Since highly correlated variables explain much of the same variation, including both will often be redundant. Meaning that when both are included, one of them will probably turn out to be non-significant in a regression. Through this section we will provide additional explanation for relationships that are considered important for the setup of our regression model. For a complete summary of the correlation matrix, please refer to Table 9, in the appendix.

Fund sequence number vs firm experience

The correlation matrix indicates a correlation coefficient of 0.6388 between fund sequence number and firm experience, and 0.7568 between LN(sequence number) and firm experience. This coefficients are considered moderate to high degree of correlation. This indicates that firm experience, measured by the number of years between the vintage year of a firm's first fund and current fund, will be explained partly by the sequence number of a fund. The intuition behind this is that as time increases from the vintage year of a PE firm's first fund, the fund sequence number also increases. They move in the same direction and thus explain much of the same variation. Because of this we believe that our instrument for detecting firm experience, which in theory is assumed to have an positive impact on the portion of IPOs of a fund, is to some extent covered by the inclusion of LN(sequence number) as an explanatory variable in our model.

Geometric average return of the MSCI world Index vs Pre2000 dummy

The correlation coefficient of 0.8378 between these two variables, necessitate further assessment. In our data, only three years after the millennium are included, but our measure of MSCI World geometric returns take use of perfect information of the index returns in the corresponding 10 year time horizon as the vintage year of the PE investment. Knowing that stock markets during this period was subject to severe financial crisis, first through the burst of the dot-com bubble in 2001, and later by the financial crisis of 2008, the direction and magnitude of the correlation is understandable. Post 2000 the returns of stock were low compared to pre 2000 returns. Thus pre 2000 observations with a dummy equal to 1 is associated with higher values of geometric return of the MSCI World index. The high degree of correlation is not seen as a problem for our analysis as it is natural feature that a dummy will have high correlation with the other variable under the above circumstances.

The inclusion of Pre2000 dummy seek to elaborate whether the IPOshare is higher for pre 2000 investments, as the IPO activity in general are believed to be lower during bear markets. If a bear market places too low values on firms the IPO becomes less attractive. For extensive arguments about Market-Timing theories, please refer to Ritter and Welch (2002).

10 year geometric average MSCI world vs US 10 year Treasury Bond yield

10-Year geometric average of the MSCI world is correlated with the 10 year bond yield of US Treasury bonds , with a coefficient of 0.7516. A strong positive linear relationship seems to exist and it is valuable to deepen this further as the economic reasoning behind it is somehow ambiguous.

US Treasury bonds are used as a proxy for the risk free rate, and matched with the duration of the investments being analysed. From economic theory, more specifically the CAPM model, we know that when the risk free rate increases, the cost of capital will increase as well, holding other factors equal. Thus an increased risk free rate will reduce the net present value of a stock following a discounted cash flow argument. This effect will undermine stock market returns.

However, with rising interest rates all other factors will indeed not be held constant. As a general rule, an increase in interest rates corresponds with significant economic growth in the economy. Therefore a period with rising interest rates also reflects that the economy is growing and that there is an increase in corporate profits. Thus, although stock prices are mitigated by the increased interest rate effect, the favourable economic growth and increase in earnings, on average, more than offset the negative effect. However we should remark that not all industry segments react to this change in the same way. Certain industries carry far higher debt levels than others. So when interest rates rise, the carrying costs of these higher debt levels may weigh more heavily on these relatively disadvantaged industries.

Fund Size vs Sequence Number

Another aspect we discovered in our preliminary studies was that the correlation between fund size and sequence number was relatively low, 0.009, as seen by Table 9. In similar empirical studies one normally observes the opposite because funds that perform well have a tendency to attract more capital to their consecutive funds. Thus an increase in sequence number of a fund should in some way be associated with a higher fund size, based on the assumption that increase in sequence numbers indicate that the firm has performed well previously. The observed correlation however might be due to differences across firms. All firms may have increasing fund size as a function of sequence number, but certain firms may have a larger initial fund size compared to funds of smaller firms.

5.2 Hypothesis I: Specialization Effects on Fund Performance

As mentioned in section 3.2, the performance of funds is approximated by three definitions of success according to the expected exit preference and related returns. Having earlier explained some of the variables considered in the regression analysis, this section analyse the results depicted in Table 3 to Table 6 in the Appendix. A full description of the variables can be located in Table 8.

5.2.1 IPO-share

Initially we ran a plain multiple regression on our set of variables, without allowing for any arbitrary form of heteroskedasticity or autocorrelation. Making conclusions solely on this regression result is without doubt not very robust, and it would most probably lead to false interpretation. As pointed out in the alternative methods section, the problem of OLS when modelling bounded fractional responses is that the real marginal effects of the regressors are non-constant throughout the unit range and that OLS fail to deal with extreme observations properly. The predicted value of the response might lie outside the defined limits. Because of this, the results of our models relying on OLS is limited to describe the effects for those funds that are located around the middle of the distribution, where a linear relationship is more probable. The problems of OLS, can also be seen by considering the residual versus fitted plots for the OLS regressions. They show distinct lines that limits the size of the residuals. Due to the argumentation above, we will not build our analysis of IPO share on the plain OLS, but rather on the two-way clustering approach, the fractional GLM model and the two-part model.

An additional aspect that needs to be clarified before the discussion of our results, is the skewed distribution of IPO shares in our sample. Approximately one third of measured IPO shares report that the funds have not exited any of their portfolio companies through an IPO. If one consider the distribution of HHI industry upon these funds, one clearly see that a large amount of the observations are located at HHI industry equal to one, i.e. full specialization. This property is mainly a direct consequence of our definition of HHI and the fact that most of these funds only have one portfolio company. The above factors leads the effect of industry specialization to decrease using the whole sample, and it seems reasonable that there are different processes determining whether a fund have zero or a fraction of their portfolio companies exited via an IPO. Based upon this, we consider the sample in two parts as suggested by Ramalho and Silva (2009). The subsample containing the continuous part of global IPO shares in PE, is considered using the second part of the two-part model, together with the two-way clustering and fractional GLM procedure. These models only assess IPO share values within the range (0, 1].

HHI Industry

Regressions (3)-(6) in Table 3 show that the effect of HHI industry is relatively small and approximately 0.08. However, the size of the coefficient is not very meaningful because we know that it is negatively biased because of the nature of our data and the definition of specialization. One can see this by considering the properties of the observations gathered at an IPO share equal to zero. 1,014 observations, or over one third of our data, is massed at this lower limit. Among these observations we find that 184 funds have only one portfolio company investment. Because of this feature, 20% of funds with zero IPO share will have an HHI industry equal to one per definition. Consequently the effect of industry specialization will be driven downwards. It makes more sense to follow the arguments of Ramalho and Silva (2009) and consider that there are two different processes underlying whether a fund has zero IPOs or positive shares of IPOs. The above argument is supported further by the results of the binary part of the two-part model, regression (1) in Table 4, which shows that the probability of having positive IPO share is reduced by 0.463 with a unit increase in HHI industry.

When considering the continuous part we observe much larger average marginal effects of industry specialization, ranging from 0.437 to 0.519. This shows that there is a very substantial effect of being more specialized with regards to industry. We would like to comment that the effect should be interpreted as an indication that highly industry specialized funds are associated with higher shares of IPOs. The magnitude of the effect should be interpreted with caution as the IPO share effect of each single IPO exit will be larger for funds with smaller number of investments in their portfolios. In addition to this, lower number of portfolio companies would be likely to increase the HHI measure, making the marginal effects observed too large. Therefore, we cannot with certainty determine the real size of the effect, only conclude that the effect is positive and located somewhere between 0.08 and 0.5. For economic purposes, this result implies that the higher the degree of specialization in industries, the higher the portion of initial public offerings should be on average.

Our results are in alignment with our specialization hypothesis and confirms that there is evidence that through greater industry expertise, the industry specialized funds picks better investments, provide additional value through extraordinary business insight or other business capabilities. If one draws the argument further, the result implies that industry specialized funds are associated with higher returns. Based on the suggestions of other research, who concludes that IPOs yields larger returns than other private equity exit strategies. We are confident regarding the validity of this result, as all of our models for IPO share shows that the HHI industry measure is significantly positive. Specifically, our result match with Gompers et al. (2005) who found evidence that industry specialized VCs tended to be more successful. It contradicts with evidence provided by Ljungqvist and Richards (2003) and Brigl et al. (2008) who concluded that industry diversification does not have significant effect on returns or the firm performance.

LN Fund size

When considering the sample as a whole, all of the included regression models in Table 3 show an estimated coefficient and average partial effect that is statistically significant, but very small. This does not mean that fund size is not important for IPOs. Looking at the binary part of the two-part model, regression (1) in Table 4, we observe an average partial effect of 0.0633. This means that the effect of a unit increase in LN Fund Size relates to an average effect of 6.3% increase in the probability of having a positive IPO share. The increase of LN fund size for small funds increases the probability of having a positive IPO share by a substantial amount, but as the fund size exceeds a certain threshold, approximately 180−200 million, the effect of increasing fund size becomes less important due to the diminishing nature of a logarithmic function. When considering the continuous regressions (2)-(5) in Table 4, we observe that there is no significant effect of an increasing fund size on the achieved IPO share for funds that already have positive shares of IPOs. Our result shows that as long as a fund has had an IPO, increasing fund size further does not affect your share of IPO exits. This result is in alignment with the intuition that only funds above a certain size have the possibility to exit an investment through an IPO.

LN Sequence Number

We find statistically significance that funds with larger sequence numbers are associated with higher share of IPOs, for regression (3)-(6) in Table 3. But the average marginal effect on IPO share between two subsequent funds is approximately zero and thus not economically significant. As indicated by the binary part of the two-part model, a unit increase of LN Sequence Number increases the probability of having a positive IPO share by 4,6%. However, the range of LN Sequence Number is limited to the interval [0, 4.1] and thus the effect is very small. As indicated by our results for the continuous part, regressions (2)-(5) in Table 4, a unit increase of LN Sequence Number does not substantially increase the IPO share for funds that have strictly positive IPO shares.

As mentioned in previous sections, we have found that there is a strong correlation between firm experience and fund sequence number. We therefore expect that an increasing sequence number would take into account some of the experience obtained by General Partners. Our results show that having larger sequence numbers have some impact in achieving positive IPO shares. This could be related to the fact that larger sequence number are associated with more experienced managers with better knowledge about the dynamics in the industry and that one is able to utilize this more efficiently in the beginning, but that the experience curve is diminishing. Our results indicates that the effect of increasing sequence numbers have very little effect on the actual amount of IPO share. Compared to other empirical research on this matter, we correspond to these studies with respect to statistical significance and the direction of the relationship. But we do not find that the increase in LN Sequence Number have an important economic effect as suggested by Kaplan and Schoar (2005) and Gottschalg et al. (2004) who concludes that firm experience has positive impact on the rate of return of PE funds.

Firm Location

Our results in both Table 3 and Table 4, suggest that funds of firms located in Europe and North America is disadvantageous and yields lower IPO shares than funds of firms located in other parts of the world. All models reports a negative effect, with an estimated coefficient and average marginal effect within the range [-0.160, -0.135] for Europe and [-0.128, -0.0943] for North America.

These results are a bit surprising. We expected the financial institutions in the western part of the world to be more developed and that the quality of the domestic institutions (stock exchanges) would be a major determinant of IPO activity. In their article, Doidge et al. (2011) show that there has been a large change in IPO activity across the globe and that the activity within the U.S. has fallen compared to the rest of the world. The decline in U.S. share of the total global amount of IPOs is explained by both extraordinary growth of IPO activity in other continents, as well as a lower activity within the U.S. With respect to these findings, our results match well. Another explanation could be that the sample selection and information gathering process done by Thomson Reuters could be insufficient with respect to the private equity industry coverage in other parts of the world. The VentureXpert data base contains a lot more detailed and extensive information from the European and American private equity industry. Therefore we believe that their cover of PE firms in other continents is less detailed and possibly only contains information from the top performers, causing a positive selection bias. The consequence of a potential positive selection bias is that PE funds in the rest of world category have higher IPOshare than those of the U.S. and Europe.

Venture Capital

When examining the whole sample of global PE funds, there is no statistically significant effect of the Venture Capital dummy on IPO share. The binary model with logit link reports that Venture Capital funds have a 13.9% higher probability of having positive IPO share. The reason for this is that Venture Capital funds have a larger average number of portfolio company investments, and thus the possibility of at least one of these having an IPO is larger. There is few important conclusions to draw from this result other than the reason discussed above. Models (2)-(5) in Table 4 on the continuous subsample report a negative average partial effect of [-0.0466, - 0.0429]. Even though these results are highly statistically significant, the size is relatively small and nor is the effect reported significant when considering the entire sample.

Macroeconomic variables

In all regressions where year dummies are excluded, we have included macroeconomic variables. These variables are believed to have an impact on the attractiveness of IPOs and thus will explain in large part the level of IPOs achieved by funds. We know that IPOs are very cyclical and varies with the situations of the public equity markets and the overall financial climate. We achieve statistically significant results, confirming that positive macroeconomic events drive IPOs. On the other hand, the macroeconomic variables of the MSCI 10-year average and the PE10 measure are based on a 10 year geometric average going forward and take use of perfect information regarding the public equities market during the normal expected life time of PE funds. Because of this, the results from these measures has little economic importance for decision purposes as they are unknown at the vintage year. The result of the 10 year bond yield on US bonds is however interesting to comment because this is the only macroeconomic measure in the analysis that does not take use of perfect information. The 10-year bond yield is measured at the end of each vintage year and approximates the risk free rate over the normal fund investment horizon. Regressions (4) and (6) in Table 3 show a positive and large correlation between the bond yield and IPO share. For the continuous part, regressions (3) and (4) show a slightly lower but still highly significant effect. Based upon these results we argue that the increased rate on long term bonds drives IPO share. When the risk free rate increases, investors require higher rates of return for longer periods because one expects that stock markets are in a positive situation and further positive price development in the stock market is expected. Thus the expectation that there will be a positive development in the stock market will make IPOs more attractive and our result indicate that the expectation is a self-fulfilling prophecy and that IPO shares increases under such circumstances.

5.2.2 MA-share

Our next success measure consists of exits through mergers and acquisitions (M&As). These results are located in Table 5. In the case of M&As the coefficient of determination, adjusted R-squared, is lower than for the other success measures used. This is due to the fact that M&As are based on agreements and negotiations, and determined in a larger degree by variables that are either unobservable or hard to quantify such as chance or network-effects. As the distribution of the MAshare is more normally distributed than the IPOshare, with most observations located between 0 and 1, thus not affected with many observations at the boundaries, we find no need for running a two-part model. After running parametric OLS, two way clustering on year and firm(3), and fractional GLM regressions with and without year dummies (4) and (5), we are left with independent variables that are statistically significant and mostly similar in direction and magnitude. In regressions where we include year dummies (YD), namely regressions (1),(2) and (4), we exclude macro economic variables as they are collinear with the time dummies (Thompson, 2010).

HHIindustry

According to our results there is no real relationship between the industry specialization and MAshare. Regressions (1) and (5) show statistically significance with p-values of 0,01 and 0.05, respectively. The size of the coefficient in (1) and the average marginal effects reported in (5) are too small to have any impact of economical importance on MAshare, thus we choose to ignore the effect of industry concentration on MAshare.

HHICountry

We find a concave relationship with the geographic concentration measure HHIcountry and MAshare reaching its maximum at roughly 0.15 across regression models when HHIcountry equals 0.77. The economic impact would be that a high degree of domestic investments increases exit rates through M&As, but only up to a certain point. This does not necessarily mean that it is the small portion of foreign investments that increases the MAshare, but rather funds not limiting themselves to only investing in one country may have a specific set of characteristics making them more successful. In addition to this it seems reasonable that if one is concentrated within a certain geographical area, one is more likely to have better business connections and a reputation that would be advantageous when negotiating with potential buyers of portfolio companies.

Firm Experience

We find a positive linear relationship with the years of experience of the GP and MAshare. But at a slope of only 0.002 or 0.2% per year of experience, this is not of any economical significance.

Active

In the case of M&As, the Active dummy is significant and negative which of course is sensible. If a fund has not been not been able to implement the restructuring of their investments yet, it will not be for sale, and the fund will have a larger fraction of active portfolio companies which in turn reduces the amount of M&As. As we already have dropped funds younger than 10 years before running our regressions, this effect is not very large and should not be emphasised.

Location of GP

One interesting observation is that both Europe and North America individually have a larger MAshare compared to the rest of the world which serves as the baseline. This could have multiple explanations, one of them being a better functioning economy which leads to an easier transaction process between participants. Another reason could be a sampling error, where American and European report all of their transactions while the rest of the world does not to the same extent. However, this would only be speculation. We find the North America-effect to be of substantial economic importance, as it has a positive effect on the MAshare of roughly 18-20 percentage points, all else equal.

5.2.3 Success = IPOshare + MAshare

In our final measure of fund success we combine the portfolio company exits in a joint measure. Using this definition, both initial public offerings and M&A transactions are defined as successful exits. The results of our regressions are located in Table 6.

HHIIndustry

Results from our regressions indicate that the industry concentration has a concave relationship on the overall success, positive throughout the allowed interval [0,1] and with a maximum at HHIIndustry = 0.67 indicating an effect on Success of 12 percentage points at maximum and 9 percentage points at the upper HHI limit. The interpretation would be that a high industry concentration is related with a higher portion of successful exits. We should note that the categorisation of industries may not be perfect, and could thus explain the non-linear relationship. Two categories may be more closely related than others, e.g. computer software and hardware are more similar than for instance fishing and electronics. This results in a hi-tech fund specializing in computer related companies, may get the same HHI-score as a generalist fund. Either way, the effect is still positive and of both statistical and economical significance.

Fundsize

In line with other researchers such as Kaplan and Schoar (2005) and Gottschalg et al. (2004), we find a statistically significant positive relationship between success and the fundsize. Our results indicate a diminishing marginal effect, and do not have any major influence until the fundsize exceeds 160$m which results in a Success increase of 1000 basis points, all else equal.

Pre2000

A large effect on success is our Pre2000-dummy, indicating that funds established before year 2000 have a larger portion of successful exits compared to those established after 2000. This is to be expected as the funds established recently may still be active and in the process of restructuring their portfolio companies, and due to the fact that most of them held portfolio companies during the financial crisis of 2008, reducing both IPO and M&A activity substantially.

5.3 Hypothesis II: Differences in the Specialization Effect on IPOshare between Venture Capital and Buyout Funds

After using both OLS and GLM approach on IPOshare in hypothesis 1, the coefficients turns out quite similar between regression methods. For simplicity reasons we will stick with the 2-way clustering and fractional GLM for hypothesis II and III. To check for differences we generated a new variable as the product of the Buyout or VentureCapital dummy and HHI industry. By doing this we can assess whether the effect of industry specialization on IPOshare differs between groups. In our analysis we do not find any statistical evidence that such a difference between the effects of specialization exist as both BOHHI and VCHHI coefficients turn out to be statistically insignificant. This indicates that effects of specialization found while testing $H1$ are not dependent on the funds' preferred investment type. We did not find it useful to include a table of non-significant variables, but the procedure and regressions are located in our Stata do-file.

5.4 Hypothesis III: The Degree of Specialization and Fund Failure Rate

By examining statistically significant coefficients in regression for both IPOshare and failure rate (located in Table 7), we find that only one variable is statistically significant in both cases and has the same direction. The interest rate for 10 year US bonds in the fund's vintage year is positively related to both an increase in IPOshare as well as the failure rate. That a higher interest rate on US Treasury bonds, which serves as a proxy for risk free rate, would lead to a higher cost of capital through the use of the capital asset pricing model (CAPM) would increase the failure rate is not surprising given the increased financial stress to portfolio companies and less willingness from investors to invest as the net present value will be reduced.

As mentioned in section 5.1.1 regarding the interest rate on 10 year US treasury bonds, this effect can also benefit the IPO share. What we can conclude from this, is that macro economic events can increase volatility in the market, and that we do not have sufficient evidence to state that funds with high IPO share also have endogenous properties associated with a large failure rate. We do however find a negative relationship between industrial specialization and failure rate. This supports the theory of strategic management theorists stating that fund specialization can be used to control portfolio risk (Norton and Tenenbaum, 1993).

5.5 Acknowledgements

Empirical studies in finance often use proxies for variables that are difficult to quantify or are in other ways unobservable. Any difference between the true variable of interest and the proxy results in a measurement error. When variables are measured imperfectly, the measurement error becomes part of the regression error. The scope of the error on coefficient estimates, depends on the errors statistical properties and measurement error does not necessarily create attenuation bias in the estimated coefficients (Roberts and Whited, 2012). This kind of measurement errors can be happening in our study, but because we use well documented proxies from other studies and only take us of information from reliable sources, we state that the potential measurement error is not severe. We strengthen this claim by observing that the estimated coefficients are fairly similar between the models using the parametric year dummy approach and those that take use of macroeconomic variables.

We also would like to state that even though we have worked hard to be as thorough as possible with the treatment of our data, we cannot guarantee the accuracy 100%. Errors may originate from the possible survivorship bias discussed in earlier sections or from other types of imperfections from the VentureXpert source at Thomson Reuters.

6 Conclusion

The aim of this paper was to provide empirical evidence to document the effect of a private equity fund's industrial or geographical specialization on its success and to investigate if these effects differ between venture capital and buyout funds. This would in turn strengthen the specialization hypothesis derived by Lossen (2006) stating that being specialized within a field of expertise increases the ability to find potentially good investments by reducing the information gap, in addition to add more value by implementing more of their know-how on a managerial level. We find that industry specialization has a positive effect on the IPO share of a fund's portfolio, and the magnitude is of great importance especially after a fund has had at least one IPO exit. We also find that one of the key endogenous factors, namely fund size, is a key driver to achieve positive shares of IPOs. We also find that geographical specialization has a positive non-linear relationship with the share of M&As. Our results also show that the Europe and North America have a lower IPO rate than rest of the world, but a higher share of trade sales. We could not find any significant differences between investment stages in regards to the specialization effect on IPOs and that the only statistically significant variables that effect both IPO share and failure rate the same way, is the interest rate of 10 year US Treasury Bonds.

By using several different econometric approaches, robust on both firm and time effects, on a large set of observations that all lead to approximately the same coefficients we strengthen the validity of our results. We have chosen to use a more indirect approach to measure success than the use of IRR. This removes the possibilities for PE-firms misreporting their returns.

Our results have several implications. If one believes that IPOs are more profitable than trade sales, LPs should invest in specialized PE funds, and GPs should focus on narrowing their investment scope to as few industries as possible, given that they have superior knowledge in this field relative to other industries. It has also an implication for entrepreneurs with a goal of going public with the financial backing of PE firms.

However, there are some limitations to this paper. First, the dataset used may be subject to survivorship bias as only surviving PE-firms report in to VentureXpert. Therefore it may be the case that also failed firms may have some of the same characteristics as the one we classify as successful, which would indicate that high success comes at the cost of high risk. Second, our results may suffer from omitted variable bias, meaning that important variables are missing from the regression. This would not affect our coefficients, unless the missing variables were to be correlated with the independent variables in our regressions.In our opinion, there are no omitted variables that are correlated with the regressors. Thus, we do not consider the omitted variable bias to influence our results. Third, we acknowledge that our results may not hold for all industries. Some industries will possibly have a higher share of IPOs compared to others. A study on PE success of specialized funds across different industries would be an interesting topic for future research.

7 Appendix

7.1 Figures

Figure 1: **Structure of Private Equity**

The general partner (GP) manages PE funds which in turn invest in several portfolio companies. The limited partners (LPs) contributes with the majority of the capital to the fund, and are thus entitled to most of the capital gains from the investments. The GP collects annual fees, on average 2% of capital under management. Figure adapted from Gilligan and Wright (2010).

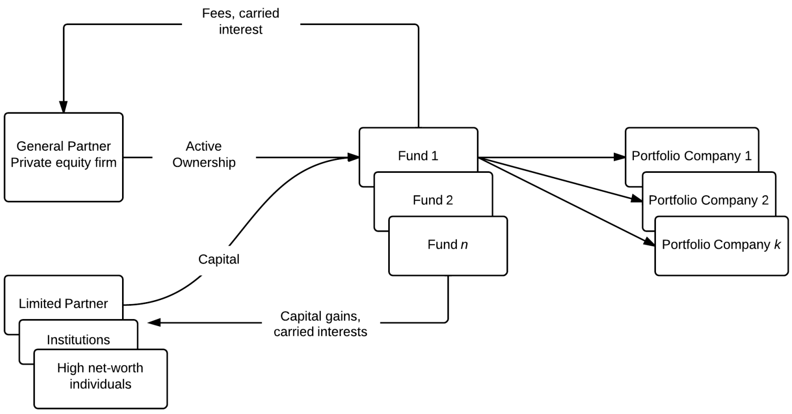
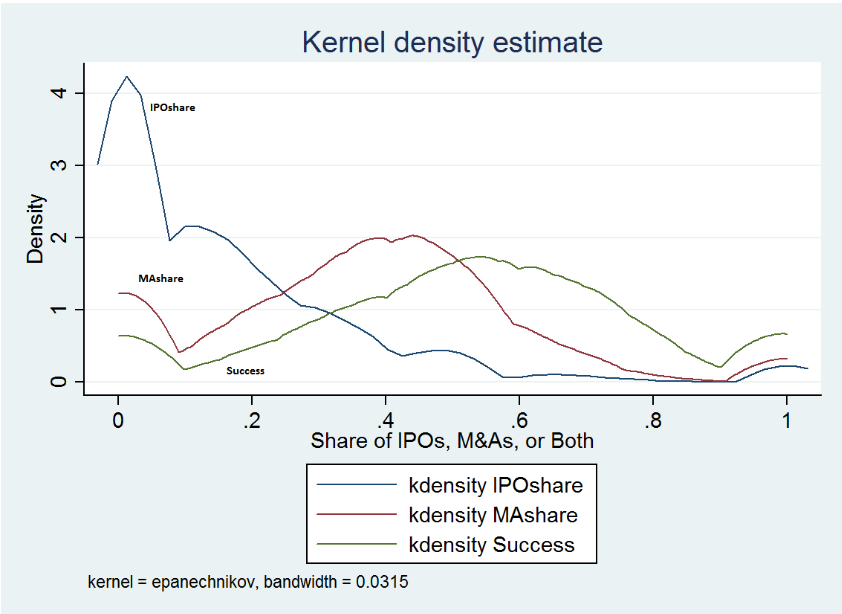


Figure 2: **Distribution of Exits**

The distribution of IPOshare, MAshare and the overall Success.



7.2 Tables

7.2.1 Composition of Sample

Table 1: **Composition of Sample**

The dataset is collected from VentureXpert, containing funds from 1960 to 2013. Due to quality concerns all funds prior to 1975 have been dropped as well as funds with a vintage year later than 2003 because they have not operated long enough to measure their lifetime performance. After all adjustments have been made, we are left with a sample containing 2,898 funds. The funds are from all investment stages, but with a majority of VC funds. 75 % of the funds are located in North America. The table below shows that the standard deviation is much larger than the mean, indicating a positive skewness of the data with many small observations and a few large outliers in terms of fundsize.

	Funds	Fund Seq Number	Fund Size				
	Obs	Mean	Mean	Std. Dev.	Min	Median	Max
All funds	2,898	4.32	158.71	548.58	0.1	50	22,887
Fund Type							
Buyout	407	4.67	367.73	628.86	0.2	125.9	4,326.50
Venture Capital	2,278	4.1	99.94	218.85	0.1	37.3	5,000
Other Private Equity	213	5.96	387.85	1,633.07	0.6	135	22,887
Continent of PE Firm							
Europe	476	4.17	247.21	1,146.82	0.1	57.55	22,887
North America	2,179	4.29	147.19	328.40	0.1	60	5,340
Rest of World	243	4.89	88.68	170.31	0.6	35	1,895
Fund Sequence Number							
1	899	1	115.22	798.46	0.2	30.9	22,887
2	555	2	119.85	264.60	0.4	42	2,643.8
3	363	3	129.76	304.66	0.1	50	4,326.5
Later	1,081	8.71	224.54	450.73	0.1	83.2	5,340
Liquidation Status							
Active	1,357	4.75	210.86	453.20	0.1	65.3	5,340
Liquidated	1,541	3.94	112.78	617.07	0.1	35	22,887

Table 2: **Fund Characteristics**

The fund size measure reflects the committed capital reported in VentureXpert and is recorded after the last funding round in the beginning of a fund's life cycle. The sequence number reflects how many funds the GP has raised up to and including the fund in question. 899 of the funds in our sample have a sequence number of 1, thus classified as First Fund. Fund success is based on fraction of a fund's portfolio companies that have exited through an Initial Public Offering(IPOshare), a trade sale (MAshare), or a combination of these two (Success rate). Failure rate depicts the fraction of a fund's portfolio company that have gone bankrupt or are in other ways defunct. GP's experience is measured in years starting at 0 for funds with sequence number = 1 and up to the vintage year of fund i. HHI, or Hirschman-Herfindahl Index, is used as a measure of investment concentration in industries or countries and has a value from 0 to 1 where 1 is perfectly concentrated and 0 is infinitely diversified.

Fund Characteristics	Obs.	Mean	Std. Dev.	Min	Median	Max
Fund Size	2,898	158.71	548.58	0.1	50	22,887
Sequence Number	2,898	4.32	5.27	1	2	61
First Fund (fraction, %)	899	31.02				
Fund Success						
IPO share (%)	2,898	15.85	20.04	0	10.00	100
M&A share (%)	2,898	36.49	23.52	0	37.50	100
Success rate (%)	2,898	57.05	26.04	0	59.09	100
Failure rate (%)	2,898	15.04	18.17	0	10.53	100
GP's experience						
Years since GP's first investment	2,898	6.02	7.54	0	3	40
No. of portfolio co. owned by GP	2,898	65.30	147.20	0	13	1483
Fund Specialization						
HHI Industry	2,898	0.43	0.26	0.09	0.35	1
HHI Geography	2,898	0.87	0.22	0.12	1	1

7.2.2 Regression Results

Over the next 5 pages you will find our regression results. Table 3 and Table 4 depicts results for the analysis of IPO share. Table 5 relates to MA share, and Table 6 to the Success rate. Table 7 relates to Hypothesis III. The variables used in our regressions are named in such a manner that the reader should be able to understand them. Nevertheless, a more thorough explanation of the variables can be located in Table 8 if needed.

Table 3: **IPOshare**

Different regression models based on IPOshare as the dependent variable. For regression (2), (3) and (5) the year dummies are included but not reported. Coefficients for the most part are highly statistically significant across models and without substantially differences in magnitude. LN(Fundsize), HHIIndustry, LN(SeqNum) MSCI World Index and PE10 are all positive reflecting that the size, industrial concentration, experience as well as macro economic factors have a positive impact on IPOshare. If the fund is located in Europe or North America it will have a negative impact on the fund's IPOshare, compared to the rest of the world that serves as the baseline. (5) and (6) are fractional GLM with and without year dummies, respectively. Numerical figures of (5) and (6) are margins dy/dx (average partial effect). Macro economic variables not included for (2),(3) and (5) due to inclusion of year dummies.

	(1) OLS	(2) OLS with YD	(3) OLS with YD and clustered by firm	(4) 2-way clustering, time and firm	(5) GLM with YD dy/dx	(6) GLM without YD dy/dx
VARIABLES	IPOshare	IPOshare	IPOshare	IPOshare	IPOshare	IPOshare
LN(FundSize)($m)	0.0131***	0.0144***	0.0144***	0.0127***	0.0148***	0.0132***
	(0.00241)	(0.00241)	(0.00346)	0.00385	(0.00345)	(0.00343)
LN(SeqNum)	0.0184***	0.0187***	0.0187***	0.0183***	0.0202***	0.0193***
	(0.00404)	(0.00402)	(0.00503)	(0.0049)	(0.00476)	(0.00484)
HHIIndustry	0.0801***	0.0852***	0.0851***	0.079***	0.0831***	0.0772***
	(0.0142)	(0.0142)	(0.0228)	(0.0235)	(0.0206)	(0.0208)
Europe	-0.141***	-0.139***	-0.139***	-0.141***	-0.135***	-0.139***
	(0.0149)	(0.0148)	(0.0220)	(0.025)	(0.0191)	(0.0191)
NorthAmerica	-0.104***	-0.106***	-0.106***	-0.104***	-0.0956***	-0.0943***
	(0.0128)	(0.0128)	(0.0206)	(0.021)	(0.0155)	(0.0155)
MSCITenYearFwdAvg	0.815***			0.6561***		0.797***
	(0.139)			(0.1272)		(0.121)
Bondyieldtenyear	1.400***			1.652***		1.294***
	(0.295)			(0.349)		(0.220)
PE10	0.255***			0.280***		0.272***
	(0.0597)			(0.126)		(0.0553)
Pre2000		0.0930***	0.0931***		0.104***	
		(0.00994)	(0.0113)		(0.0122)	
spreaddummy	-0.0335**					
	(0.0141)					
SpreadBondBill	1.351***					
	(0.489)					
Constant	-0.0301	0.0784***	0.0786***	-0.057		
	(0.0318)	(0.0186)	(0.0290)	(0.0583)		
Observations	2,898	2,898	2,898	2,898	2,898	2,898
Adj. R-squared	0.126	0.138	0.138	0.124		

Standard errors in parentheses

$***p < 0.01, **p < 0.05, *p < 0.1$

Table 4: **2-part and fractional model**

(1) and (2) represent the two parts of a two part fractional model, where (1) is a binary model with logit link for IPOshare = 0, and IPOshare > 0. The second part (2) is continuous in for the interval (0.1]. (3) is a 2-way clustered OLS clustered on Firm and Year for all observations except for IPOshare = 0. Year dummies for (1), (2) and (5) are included but not reported. We find that when we exclude IPOshare = 0 the positive relationship between industry specialization and IPOshare increases substantially. This is most likely due to the inherent effect where funds with only 1 company automatically recieve a HHI value of 1, and with large likelihood IPOshare = 0, reducing the real connection between the variables. See Table 8 for descripton of variables.

	(1)	(2)	(3)	(4)	(5)
	Two-part model				
	1.part binary, logit link dy/dx	2.part frac, logit link dy/dx	2-way clustering time and firm (0,1]	GLM without YD dy/dx(0,1]	GLM with YD dy/dx(0,1]
VARIABLES	IPOshare	IPOshare	IPOshare	IPOshare	IPOshare
LN(FundSize)	0.0633***				
	(0.00664)				
HHIcountry	-0.206***				
	(0.0484)				
HHIIndustry	-0.463***	0.442***	0.519***	0.437***	0.442***
	(0.0302)	(0.0234)	(0.0281)	(0.0232)	(0.0234)
LN(SeqNum)	0.0456***	0.0199***	0.0210***	0.0182***	0.0199***
	(0.0115)	(0.00526)	(0.00524)	(0.00530)	(0.00526)
Europe	-0.0654**	-0.148***	-0.160***	-0.151***	-0.148***
	(0.0281)	(0.0210)	(0.0244)	(0.0213)	(0.0210)
NorthAmerica		-0.116***	-0.128***	-0.114***	-0.116***
		(0.0188)	(0.0207)	(0.0191)	(0.0188)
VentureCapital	0.139***	-0.0438***	-0.0466***	-0.0429***	-0.0438***
	(0.0217)	(0.0129)	(0.0126)	(0.0127)	(0.0129)
MSCITenYearFwdAvg			0.514***	0.587***	
			(0.182)	(0.127)	
Bondyieldtenyear			1.563***	1.316***	
			(0.414)	(0.222)	
PE10			0.335***	0.304***	
			(0.110)	(0.0553)	
Pre2000	2.847***				
	(0.119)				
Contant			-0.0295		
			(0.0519)		
Observations	2,898	1,884	1,884	1,884	1,884
R-squared			0.390		

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 5: **MAshare**

This table shows different regression models based on the portion of portfolio companies that was exited through a M&A as the dependent variable. For regressions that include year dummies (YD), namely (1),(2) and (4), we have excluded macro economic variables due to collinearity. The year dummies for said regressions are included but not reported. See table 8 for descripton of variables.

	(1) OLS with YD	(2) OLS with YD, clustered by firm	(3) 2-way cluster year and firm	(4) GLM with YD dy/dx	(5) GLM without YD dy/dx
VARIABLES	MAshare	MAshare	MAshare	MAshare	MAshare
HHIIndustrySqr	-0.0450***				-0.0524**
	(0.0147)				(0.0262)
HHIcountry	0.381***	0.394**	0.396***	0.426**	0.401***
	(0.146)	(0.132)	(0.113)	(0.143)	(0.136)
HHIcountrySqr	-0.246**	-0.264***	-0.269***	-0.280***	-0.254***
	(0.104)	(0.0954)	(0.0826)	(0.101)	(0.0961)
FirmExperience	0.00233***	0.00223***	0.00222***	0.00214***	0.00203***
	(0.000562)	(0.000663)	(0.0007)	(0.000634)	(0.000597)
MSCITenYearFwdAvg					-0.372**
					(0.148)
Active	-0.0257***	-0.0280**	-0.0304**	-0.0264**	-0.0287***
	(0.00871)	(0.0117)	(0.0144)	(0.0114)	(0.0106)
Pre2000		0.215***		0.200***	0.0431***
		(0.0218)		(0.0221)	(0.0165)
Buyout	0.0737***	0.0787***	0.0779***	0.0798***	0.0723***
	(0.0193)	(0.0241)	(0.0250)	(0.0265)	(0.0260)
VentureCapital	0.0808***	0.0878***	0.0885***	0.0899***	0.0831***
	(0.0163)	(0.0205)	(0.0214)	(0.0228)	(0.0222)
Europe	0.0948***	0.103***	0.0994***	0.117***	0.114***
	(0.0180)	(0.0239)	(0.0209)	(0.0301)	(0.0297)
NorthAmerica	0.173***	0.185***	0.182***	0.199***	0.194***
	(0.0160)	(0.0203)	(0.0204)	(0.0258)	(0.0254)
Constant	0.0234	-0.0203	0.00359		
	(0.0511)	(0.0478)	(0.0380)		
Observations	2,898	2,898	2,898	2,898	2,898
Adj. R-squared	0.086	0.084	0.079		

Standard errors in parentheses

$***p < 0.01, **p < 0.05, *p < 0.1$

Table 6: **Success = IPOshare + MAshare**

This tables shows different regression models based on the overall rate of successful exits, in other words the combined share of IPOs and M&As. For (1), (2) and (4) time dummies are included but not reported.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	OLS with TD Success	OLS with TD, cluster by firm Success	2-way cluster, time and firm Success	GLM with YD Success	GLM without YD Success
LN(FundSize)($mil)	0.0204***	0.0207***	0.0193***	0.0206***	0.0196***
	(0.00336)	(0.00405)	(0.00441)	(0.00404)	(0.00407)
LN(SeqNum)	0.0164**				
	(0.00784)				
Active	-0.0293***	-0.0330***	-0.0447***	-0.0323**	-0.0393***
	(0.0109)	(0.0127)	(0.0164)	(0.0125)	(0.0122)
HHIIndustry	0.349***	0.342***	0.288***	0.342***	0.300***
	(0.0837)	(0.0890)	(0.0857)	(0.0883)	(0.0876)
HHIIndustrySqr	-0.261***	-0.251***	-0.209**	-0.252***	-0.221**
	(0.0714)	(0.0886)	(0.0918)	(0.0873)	(0.0873)
Europe	-0.0447**				
	(0.0196)				
NorthAmerica	0.0716***	0.0983***	0.0999***	0.0963***	0.0991***
	(0.0168)	(0.0153)	(0.0143)	(0.0150)	(0.0153)
VentureCapital	0.0969***	0.0980***	0.0993***	0.0968***	0.0987***
	(0.0181)	(0.0240)	(0.0261)	(0.0247)	(0.0243)
Buyout	0.0755***	0.0739***	0.0771***	0.0723**	0.0730***
	(0.0209)	(0.0275)	(0.0229)	(0.0281)	(0.0279)
Bondyieldtenyear			1.662***		1.780***
			(0.452)		(0.312)
MSCITenYearFwdAvg			0.685***		0.616***
			(0.144)		(0.145)
FirmExperience	0.00258***	0.00396***	0.00384***	0.00390***	0.00380***
	(0.000929)	(0.000757)	(0.00094)	(0.000767)	(0.000761)
PE10					0.237***
					(0.0870)
Pre2000	0.208***	0.201***		0.200***	
	(0.0170)	(0.0168)		(0.0167)	
Constant	0.136***	0.120***	0.0613		
	(0.0386)	(0.0393)	(0.0531)		
Observations	2,898	2,898	2,898	2,898	2,898
Adj. R-squared	0.147	0.145	0.131		

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 7: **IPOshare Vs. Failure Rate**

This table shows a comparison of statistically significant variables for an OLS regression clustered both on time and firm dimensions. (1) uses IPOshare as the dependent variable, while (2)-(4) consist of the share of failed portfolio companies, either by bankruptcy or companies that are in other ways defunct. Year dummies are included but not reported for regression (4). We find that only the interest rate of 10 year US treasury bonds at the vintage year is significant with the same direction for IPO share and Failure rate.

	(1) 2-way cluster, time and firm	(2) 2-way cluster, time and firm	(3) GLM	(4) GLM with YD
VARIABLES	IPOshare	Failure rate	Failure rate	Failure rate
LN(FundSize)($mil)	0.0127***	-0.00595**	-0.00757***	-0.00947***
	(0.00385)	(0.00254)	(0.00294)	(0.00286)
LN(SeqNum)	0.0183***			
	(0.0049)			
Active		-0.0318***	-0.0226***	-0.0227***
		(0.00904)	(0.00831)	(0.00797)
HHIcountry		-0.154**		
		(0.0768)		
HHIcountrySqr		0.124**		
		(0.0532)		
HHIIndustry	0.079***	-0.0493***	-0.0587***	-0.0633***
	(0.0235)	(0.0152)	(0.0220)	(0.0216)
SeqNum		-0.00146**	-0.00255***	-0.00242**
		(0.000616)	(0.000953)	(0.000982)
Europe	-0.141***			
	(0.0247)			
NorthAmerica	-0.104***	0.0664***	0.0881***	0.0891***
	(0.021)	(0.0132)	(0.0108)	(0.0106)
Buyout		-0.0401***	-0.0401***	-0.0370***
		(0.0129)	(0.0139)	(0.0138)
MSCITenYearFwdAvg	0.6561***			
	(0.1272)			
Bondyieldtenyear	1.652***	0.991***	0.416***	
	(0.349)	(0.244)	(0.148)	
Pre2000		0.0599***	0.0969***	0.271***
		(0.0210)	(0.0114)	(0.0246)
PE10	0.280***			
	(0.126)			
Spreaddummy			-0.0183**	
			(0.00827)	
Constant	-0.0566	0.0948***		
	(0.0584)	(0.0348)		
Observations	2,898	2,898	2,898	2,898
Adj. R-squared	0.124	0.162		

Standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 8: **Description of Variables**

Dependent Variables	
IPOshare	The portion of a fund's portfolio that went public.
MAshare	The portion of a fund's portiolio that was involved in a merger or acquisition.
Success	The Combination of both IPOshare and MAshare.
Failure rate	The portion of a fund's portfolio that went bankrupt or in other ways defunkt.

Independent Variables	
HHIIndustry	Measurement of industry concentration. 1=Perfectly concentrated, 0=perfectly diversified.
HHIcountry	Measurement of geographic concentration. 1=Perfectly concentrated, 0=perfectly diversified.
HHINaive	Measurement of capital allocation or investment concentration in portfolio companies. 1=Perfectly concentrated, 0=perfectly diversified.
FundSizeMil	The capital inflow to the fund, measured in MUSD at the beginning of a fund's life cycle
SeqNum	The sequence number of each PE fund.
Fundinvestmenttype	The preferred investment stage of a PE-fund. VC, BO or Other
FundVintageYear	The year in which the fund was established
FirmExperience	Years of PE-firm experience at the start of each fund.
Bondyieldtenyear	Yearly interest rate on 10-year US bond at each vintage year.
yearTbillyield	Yearly interest rate on 1-year US Treasury bills at each vintage year.
SpreadBondBill	The difference in interest rates between 10 year bond and 1 year T-bills.
MSCIIndex	The index value of MSCI - Global
MSCITenYearFwdAvg	The geometric average change in MSCI global over a 10 year period starting in each vintage year going forward.
Active	Dummy that equals 1 if fund is currently active, 0 if liquidated.
FirmContinent	The continent in which the GP is located. Europe, N. America or Rest of the world.
GDPgrowth	The global GDP growth at each vintage year.
Pre2000	Dummy that equals 1 if vintage year is prior to 2000, 0 otherwise.
PE	The Price-Earnings ratio for each vintage year.
PE10	Geometric average of Schiller's price-earnings ratio based on inflation-adjusted earnings for a 10-year time period starting in the vintage year.

Table 9: **Cross-correlation table**

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
IPOshare(1)	1.000										
Fund Size($ Mil)(2)	0.009	1.000									
LN(FundSize)(3)	-0.026	0.438	1.000								
HHI-Industry(4)	0.063	-0.078	-0.177	1.000							
HHIIndustrySqr(5)	0.068	-0.057	-0.157	0.975	1.000						
HHI-country(6)	0.005	-0.196	-0.259	0.162	0.166	1.000					
HHIcountrySqr(7)	0.010	-0.190	-0.260	0.165	0.172	0.991	1.000				
Seq.Num(8)	0.026	0.109	0.191	0.058	0.079	-0.111	-0.104	1.000			
LN(SeqNum)(9)	0.029	0.107	0.234	0.020	0.042	-0.112	-0.106	0.847	1.000		
MSCI Ten Year Fwd Avg.(10)	0.252	-0.111	-0.329	-0.082	-0.060	0.143	0.148	-0.181	-0.201	1.000	
Firm Experience(11)	0.022	0.111	0.214	-0.031	-0.008	-0.014	-0.016	0.604	0.750	-0.142	1.000
Bond yield ten year(12)	0.241	-0.092	-0.329	-0.129	-0.100	0.147	0.151	-0.177	-0.212	0.752	-0.159
spreaddummy(13)	0.026	-0.055	-0.043	0.034	0.034	0.017	0.021	-0.025	-0.022	0.332	-0.009
PE10(14)	0.121	0.010	-0.071	0.034	0.044	0.013	0.015	-0.081	-0.054	0.187	-0.040
Active(15)	-0.157	0.089	0.216	-0.062	-0.062	-0.119	-0.121	0.077	0.088	-0.472	0.087
Pre2000(16)	0.192	-0.088	-0.220	-0.014	-0.001	0.088	0.092	-0.149	-0.146	0.839	-0.092
Buyout(17)	0.003	0.154	0.258	0.018	0.016	-0.159	-0.155	0.027	0.021	-0.018	-0.040
VentureCapital(18)	0.003	-0.205	-0.326	-0.077	-0.085	0.133	0.126	-0.079	-0.065	0.058	0.004
Europe(19)	-0.107	0.072	0.045	-0.013	-0.011	-0.344	-0.344	-0.013	0.028	-0.073	-0.047
NorthAmerica(20)	0.013	-0.037	-0.006	-0.043	-0.042	0.414	0.420	-0.010	-0.020	0.114	0.089
Spread Bond - Bill(21)	0.029	-0.032	-0.039	0.022	0.022	0.005	0.009	-0.004	-0.013	0.091	-0.014

Variables	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
Bond yield ten year(12)	1.000									
spreaddummy(13)	-0.095	1.000								
PE10(14)	0.096	0.061	1.000							
Active(15)	-0.391	-0.098	-0.196	1.000						
Pre2000(16)	0.459	0.448	0.251	-0.478	1.000					
Buyout(17)	-0.070	0.061	0.069	-0.015	0.025	1.000				
VentureCapital(18)	0.099	-0.066	-0.053	-0.013	0.015	-0.775	1.000			
Europe(19)	-0.086	-0.005	0.033	0.129	-0.056	0.204	-0.182	1.000		
NorthAmerica(20)	0.137	0.011	-0.023	-0.100	0.064	-0.173	0.135	-0.772	1.000	
Spread Bond - Bill(21)	-0.101	0.617	0.113	0.039	-0.055	0.049	-0.059	0.016	-0.009	1.000

7.3 VBA-codes

```
Sub renameboston()
strPromt = "This macro sorts out the problem with the two funds written in similar
    manner." & vbNewLine & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set ws = Worksheets("SDC-output")
    row = 2
    Do Until ws.Cells(row, "A") = ""
        If ws.Cells(row, "J") = "Bancboston Ventures" Then
            ws.Cells(row, "J").Value = "Bancboston Ventures 2"
        End If
        row = row + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub CellSplitting()
strPromt = "This macro splits the raw data from SDC-platinum into separate rows and
    columns. By running this macro compared to the 'text to column' function in excel
    all rows will be perfectly aligned with the correct information given." & vbNewLine
     & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Application.DisplayAlerts = False
    On Error Resume Next
    ThisWorkbook.Sheets("SDC-split").Delete
    On Error GoTo 0
    Application.DisplayAlerts = True
    Dim ws As Worksheet
    Set ws = Sheets.Add
    ws.Name = "SDC-split"
    Application.ScreenUpdating = False
    Set inn = Worksheets("SDC-output")
    Set out = Worksheets("SDC-split")
    Dim lastrow As Long
    Dim lastcol As Long
    Dim outlastrow
    Dim outrow As Long
    Dim row As Integer
    Dim incol As Integer
    Dim info
    Dim d
    outlastrow = out.Cells(out.Rows.Count, "J").End(xlUp).row
    lastrow = inn.Cells(inn.Rows.Count, "A").End(xlUp).row
    If inn.Cells(lastrow, "A").Value <> "end" Then
        inn.Cells(lastrow + 2, "A").Value = "end"
    End If
    lastcol = inn.Cells(1, inn.Columns.Count).End(xlToLeft).Column
    outrow = 2
    row = 2
    Size = 0
    incol = 1
    outcol = 1
    Do Until inn.Cells(row, "A") = "end"
        info = Split(inn.Cells(row, "R"), Chr(10))
        For Each d In info
```

```
            For i = 1 To lastcol
                gal = Split(inn.Cells(row, i), Chr(10))
                For Each B In gal
                    galcount = galcount + 1
                Next B
                If galcount > 1 Then
                    out.Cells(outrow, i).Value = gal(Size)
                    Else
                    out.Cells(outrow, i).Value = inn.Cells(row, i).Value
                End If
                galcount = 0
            Next i
            Size = Size + 1
            outrow = outrow + 1
        Next d
        Size = 0
tom:
        If inn.Cells(row, incol) = "" Then
            out.Cells(outrow, outcol).Value = ""
        End If
        row = row + 1
    Loop
        out.Range("A:AH").Columns.AutoFit
        Sheets("SDC-output").Select
        Range("A1:AL1").Select
        Selection.Copy
        Sheets("SDC-split").Select
        Range("A1").Select
        ActiveSheet.Paste
        Application.ScreenUpdating = True

        Sheets("SDC-split").Select
        ActiveWorkbook.Worksheets("SDC-split").Sort.SortFields.Clear
        ActiveWorkbook.Worksheets("SDC-split").Sort.SortFields.Add Key:=Range("J2:J" &
            outlastrow), SortOn:=xlSortOnValues, Order:=xlAscending, DataOption:=
            xlSortNormal
        With ActiveWorkbook.Worksheets("SDC-split").Sort
            .SetRange Range("A1:AL" & outlastrow)
            .Header = xlYes
            .MatchCase = False
            .Orientation = xlTopToBottom
            .SortMethod = xlPinYin
            .Apply
        End With


End If
MsgBox "Done"
End Sub

Sub remove_non_info_funds()
strPromt = "This macro deletes all rows that contain funds without useful information"
    & vbNewLine & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Dim lastrow As Long
    Set ws = Worksheets("SDC-split")
    For i = 1 To 2
        row = 2
```

```
        Do Until ws.Cells(row, "A") = ""
            If ws.Cells(row, "Q") = "" And ws.Cells(row, "AC") = "" And ws.Cells(row, "
                AB") = "" Then
                Do While ws.Cells(row, "Q") = "" And ws.Cells(row, "A") <> ""
                    ws.Rows(row).Delete
                Loop
            End If
            If ws.Cells(row, "O") = "" Then
                ws.Rows(row).Delete
            End If
            row = row + 1
        Loop
    Next i
End If
MsgBox "Done"
End Sub

Sub fill_in_average()
strPromt = "This macro fills in the average amount invested by the fund to portfolio
    companies where the known invested amount is missing." & vbNewLine & vbNewLine & "
    Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Dim row As Long
    Set ws = Worksheets("SDC-split")
    row = 2
    Do Until ws.Cells(row, "A") = ""
        If ws.Cells(row, "AC") = "" Then
            ws.Cells(row, "AC").Value = ws.Cells(row, "Q")
        End If
        row = row + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub delete_zerosizefunds()
strPromt = "This macros delete all funds with a fund size of zero." & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set ws = Worksheets("SDC-split")
    Dim row As Long
    row = 2
    Do Until ws.Cells(row, "A") = ""
        If ws.Cells(row, "O") = "" Then
            Do Until ws.Cells(row, "O") <> ""
                ws.Rows(row).Delete
            Loop
        End If
        row = row + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub NaiveHHI()
strPromt = "This macro is the first of five macros required to calculate the Naive HHI.
     We have deliberately split them into five subroutine macros to prevent overload of
```

```
     slower computers." & vbNewLine & vbNewLine & "The same goes for the industrial and
     geographical HHI as well" & vbNewLine & vbNewLine & "PS: This is one of the most
    demanding macros and may take a couple of minutes to complete" & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'The first of five macros required to calculate the Naive HHI. We have deliberately
         split them into five subroutine macros to prevent overload of slower computers
        .
    'The same goes for the industrial and the geographical HHI.
    Application.DisplayAlerts = False
    On Error Resume Next
    ThisWorkbook.Sheets("NaiveHHI").Delete
    On Error GoTo 0
    Application.DisplayAlerts = True
    Dim ss As Worksheet
    Set ss = Sheets.Add
    ss.Name = "NaiveHHI"
    Application.ScreenUpdating = False
    Dim p As Variant
    Dim c As Long
    Set ws = ThisWorkbook.Sheets("SDC-split")
    Dim lastrow As Long
    lastrow = ws.Cells(ws.Rows.Count, "A").End(xlUp).row
    ws.Cells(lastrow + 2, "A").Value = "end"
    Set Rng = ws.Range(ws.Range("J1"), ws.Range("J" & Rows.Count).End(xlUp))
    Set Dic = CreateObject("Scripting.Dictionary")
    Dic.CompareMode = 1
    For Each Dn In Rng
        If Not Dic.exists(Dn.Value) Then
            Set Dic(Dn.Value) = CreateObject("Scripting.Dictionary")
        End If
        If Not Dic(Dn.Value).exists(Dn.Offset(, 8).Value) Then
                Dic(Dn.Value).Add (Dn.Offset(, 8).Value), Dn.Offset(, 19)
            Else
                Dic(Dn.Value).Item(Dn.Offset(, 8).Value) = Dic(Dn.Value).Item(Dn.Offset
                    (, 8).Value) + Dn.Offset(, 19)
        End If
    Next Dn
    Set out = ThisWorkbook.Sheets("NaiveHHI")
    c = 1
    out.Range("B1") = "Output"
    For Each F In Dic.keys
        out.Cells(c, "A") = F
            For Each p In Dic(F)
                out.Cells(c, "A") = F
                out.Cells(c, "B") = p
                out.Cells(c, "C") = Dic(F).Item(p)
                c = c + 1
            Next p
            If Not F = "Fund" Then
                c = c + 1
            End If
    Next F
    Application.ScreenUpdating = True
End If
MsgBox "Done"
End Sub
```

```
Sub MakeNaiveSum()
strPromt = "This macro is step 2 of 5 of calculating the naive HHI" & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'step 2 of 5
    Set ws = Worksheets("NaiveHHI")
    Dim lastrow As Long
    Dim row As Long
    ws.Cells(1, "D").Value = "Total Sum Invested"
    ws.Cells(1, "E").Value = "Squared Investment Share"
    ws.Cells(1, "F").Value = "NaiveHHI"
    lastrow = ws.Cells(ws.Rows.Count, "A").End(xlUp).row
    If ws.Cells(lastrow, "A") <> "end" Then
        ws.Cells(lastrow + 2, "A").Value = "end"
    End If
    row = 2
    Do Until ws.Cells(row, 1) = "end"
        If ws.Cells(row, 2) = "" Then
                ws.Cells(row - 1, 4).Value = Sum
                row = row + 1
                Sum = 0
            Else
                Sum = ws.Cells(row, 3).Value + Sum
                row = row + 1
        End If
    Loop
    fundrange = ws.Range(ws.Range("A2"), ws.Range("A" & Rows.Count).End(xlUp))
    ws.Cells(1, "D").Value = "end"
    row = ws.Cells(ws.Rows.Count, "D").End(xlUp).row + 3
    Do Until ws.Cells(row, "D") = "end"
        If ws.Cells(row + 1, "D").Value = "" Then
            x = ws.Cells(row, "D")
            row = row - 1
        End If
        Do While ws.Cells(row, "A") <> "" And ws.Cells(row, "D") <> "end"
            ws.Cells(row + 1, "D").Value = x
            row = row - 1
        Loop
        ws.Cells(row + 1, "D").Value = x
        If ws.Cells(row, "D") <> "end" Then
            row = row - 1
        End If
    Loop
    ws.Cells(1, "D").Value = "Total Sum Invested"
    'Application.ScreenUpdating = True
End If
MsgBox "Done"
End Sub


Sub NaiveshareSquared()
strPromt = "This macro is step 3 of 5 of calculating the naive HHI" & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'step 3 of 5
    Set ws = Worksheets("NaiveHHI")
    fundrange = ws.Range(ws.Range("A2"), ws.Range("A" & Rows.Count).End(xlUp))
```

```
    row = 2
    Do Until ws.Cells(row, "A") = "end"
        If ws.Cells(row, "D") = "" Or ws.Cells(row, "D") = 0 Then
                row = row + 1
            Else
                ws.Cells(row, "E").Value = (ws.Cells(row, "C") / ws.Cells(row, "D")) ^
                    2
                row = row + 1
        End If
    Loop
End If
MsgBox "Done"
End Sub

Sub HHINaive()
strPromt = "This macro is step 4 of 5 of calculating the naive HHI" & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'step 4 of 5
    'Application.ScreenUpdating = False
    'Application.DisplayAlerts = False
    Set ws = Worksheets("NaiveHHI")
    fundrange = ws.Range(ws.Range("A2"), ws.Range("A" & Rows.Count).End(xlUp))
    row = 2
    For Each fund In fundrange
        If ws.Cells(row, "A") = "" And ws.Cells(row + 1, "A") = "" Then
            GoTo finish
        End If
        s = 0
        s = ws.Cells(row, "E") + s
        Do While ws.Cells(row + 1, "A") = ws.Cells(row, "A")
            row = row + 1
            s = ws.Cells(row, "E") + s
        Loop
        row = row + 1
        ws.Cells(row - 1, "F").Value = s
        If ws.Cells(row - 1, "F") = 0 Then
            ws.Cells(row - 1, "F").Delete
        End If
    Next fund
    'Application.ScreenUpdating = True
    'Application.DisplayAlerts = True
End If
finish:
MsgBox "Done"
End Sub
Sub copyNaiveHHI()
strPromt = "This macro is step 5 of 5 of calculating the naive HHI" & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'step 5 of 5
    Application.ScreenUpdating = False
    Set inn = Worksheets("NaiveHHI")
    Set out = Worksheets("SDC-split")
    Dim row As Long
    Dim rekke As Long
    row = 2
```

```
    rekke = 2
    out.Cells(1, "AS").Value = "NaiveHHI"
    Do Until out.Cells(row, "A") = "end"
        Do Until inn.Cells(rekke + 1, "A") = ""
            rekke = rekke + 1
        Loop
        Do While out.Cells(row, "J") = inn.Cells(rekke, "A") And out.Cells(row, "A") <>
             "end"
            out.Cells(row, "AS").Value = inn.Cells(rekke, "F")
            row = row + 1
        Loop
        rekke = rekke + 1
    Loop
End If
'Application.ScreenUpdating = True
MsgBox "Done"
End Sub
Sub industry()
strPromt = "This macro is step 1 of 5 of calculating the Industry HHI" & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Application.DisplayAlerts = False
    On Error Resume Next
    ThisWorkbook.Sheets("industry-split").Delete
    On Error GoTo 0
    Application.DisplayAlerts = True
    Dim ss As Worksheet
    Set ss = Sheets.Add
    ss.Name = "industry-split"
    Application.ScreenUpdating = False
    Dim p As Variant
    Dim c As Long
    Set ws = ThisWorkbook.Sheets("SDC-split")
    Dim lastrow As Long
    lastrow = ws.Cells(ws.Rows.Count, "A").End(xlUp).row
    ws.Cells(lastrow + 2, "A").Value = "end"
    Set Rng = ws.Range(ws.Range("J1"), ws.Range("J" & Rows.Count).End(xlUp))
    Set Dic = CreateObject("Scripting.Dictionary")
    Dic.CompareMode = 1
    For Each Dn In Rng
        If Not Dic.exists(Dn.Value) Then
            Set Dic(Dn.Value) = CreateObject("Scripting.Dictionary")
        End If
        If Not Dic(Dn.Value).exists(Dn.Offset(, 21).Value) Then
                Dic(Dn.Value).Add (Dn.Offset(, 21).Value), Dn.Offset(, 19)
            Else
                Dic(Dn.Value).Item(Dn.Offset(, 21).Value) = Dic(Dn.Value).Item(Dn.
                    Offset(, 21).Value) + Dn.Offset(, 19)
        End If
    Next Dn
    Set out = ThisWorkbook.Sheets("industry-split")
    c = 1
    out.Range("B1") = "Output"
    For Each F In Dic.keys
        out.Cells(c, "A") = F
        For Each p In Dic(F)
            out.Cells(c, "A") = F
            out.Cells(c, "B") = p
```

```
            out.Cells(c, "C") = Dic(F).Item(p)
            c = c + 1
        Next p
        If Not F = "Fund" Then
            c = c + 1
        End If
    Next F
    Application.ScreenUpdating = True
End If
MsgBox "Done"
End Sub

Sub MakeSum()
strPromt = "This macro is step 2 of 5 of calculating the Industry HHI" & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set ws = Worksheets("industry-split")
    Dim lastrow As Long
    Dim row As Long
    ws.Cells(1, "D").Value = "Total Sum Invested"
    ws.Cells(1, "E").Value = "Squared Investment Share"
    ws.Cells(1, "F").Value = "HHI"
    lastrow = ws.Cells(ws.Rows.Count, "A").End(xlUp).row
    If ws.Cells(lastrow, "A") <> "end" Then
        ws.Cells(lastrow + 2, "A").Value = "end"
    End If
    row = 2
    Do Until ws.Cells(row, 1) = "end"
        If ws.Cells(row, 2) = "" Then
                ws.Cells(row - 1, 4).Value = Sum
                row = row + 1
                Sum = 0
            Else
                Sum = ws.Cells(row, 3).Value + Sum
                row = row + 1
        End If
    Loop
    fundrange = ws.Range(ws.Range("A2"), ws.Range("A" & Rows.Count).End(xlUp))
    ws.Cells(1, "D").Value = "end"
    row = ws.Cells(ws.Rows.Count, "D").End(xlUp).row + 3
    Do Until ws.Cells(row, "D") = "end"
       If ws.Cells(row + 1, "D").Value = "" Then
            x = ws.Cells(row, "D")
            row = row - 1
        End If
        Do While ws.Cells(row, "A") <> "" And ws.Cells(row, "D") <> "end"
            ws.Cells(row + 1, "D").Value = x
            row = row - 1
        Loop
        ws.Cells(row + 1, "D").Value = x
        If ws.Cells(row, "D") <> "end" Then
            row = row - 1
        End If
    Loop
    ws.Cells(1, "D").Value = "Total Sum Invested"
    'Application.ScreenUpdating = True
End If
MsgBox "Done"
```

```
End Sub

Sub shareSquared()
strPromt = "This macro is step 3 of 5 of calculating the Industry HHI" & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set ws = Worksheets("industry-split")
    fundrange = ws.Range(ws.Range("A2"), ws.Range("A" & Rows.Count).End(xlUp))
    row = 2
    Do Until ws.Cells(row, "A") = "end"
        If ws.Cells(row, "D") = "" Or ws.Cells(row, "D") = 0 Then
                row = row + 1
            Else
                ws.Cells(row, "E").Value = (ws.Cells(row, "C") / ws.Cells(row, "D")) ^
                    2
                row = row + 1
        End If
    Loop
End If
MsgBox "Done"
End Sub

Sub HHI()
strPromt = "This macro is step 4 of 5 of calculating the Industry HHI" & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'Application.ScreenUpdating = False
    'Application.DisplayAlerts = False
    Set ws = Worksheets("industry-split")
    fundrange = ws.Range(ws.Range("A2"), ws.Range("A" & Rows.Count).End(xlUp))
    row = 2
    For Each fund In fundrange
        If ws.Cells(row, "A") = "" And ws.Cells(row + 1, "A") = "" Then
            GoTo finish
        End If
        s = 0
        s = ws.Cells(row, "E") + s
        Do While ws.Cells(row + 1, "A") = ws.Cells(row, "A")
            row = row + 1
            s = ws.Cells(row, "E") + s
        Loop
        row = row + 1
        ws.Cells(row - 1, "F").Value = s
        If ws.Cells(row - 1, "F") = 0 Then
            ws.Cells(row - 1, "F").Delete
        End If
    Next fund
    'Application.ScreenUpdating = True
    'Application.DisplayAlerts = True
End If
finish:
MsgBox "Done"
End Sub

Sub copyHHI()
strPromt = "This macro is step 5 of 5 of calculating the Industry HHI" & vbNewLine &
    vbNewLine & "Run this macro?"
```

```
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Application.ScreenUpdating = False
    Set inn = Worksheets("industry-split")
    Set out = Worksheets("SDC-split")
    Dim row As Long
    Dim rekke As Long
    row = 2
    rekke = 2
    out.Cells(1, "AO").Value = "HHIIndustry"
    Do Until out.Cells(row, "A") = "end"
        Do Until inn.Cells(rekke + 1, "A") = ""
            rekke = rekke + 1
        Loop
        Do While out.Cells(row, "J") = inn.Cells(rekke, "A") And out.Cells(row, "A") <>
             "end"
            out.Cells(row, "AO").Value = inn.Cells(rekke, "F")
            row = row + 1
        Loop
        rekke = rekke + 1
    Loop
End If
'Application.ScreenUpdating = True
MsgBox "Done"
End Sub

Sub Country()
strPromt = "This macro is step 1 of 5 of calculating the Geographical HHI" & vbNewLine
    & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Application.DisplayAlerts = False
    On Error Resume Next
    ThisWorkbook.Sheets("Country-split").Delete
    On Error GoTo 0
    Application.DisplayAlerts = True
    Dim ss As Worksheet
    Set ss = Sheets.Add
    ss.Name = "Country-split"
    Application.ScreenUpdating = False
    Dim p As Variant
    Dim c As Long
    Set ws = ThisWorkbook.Sheets("SDC-split")
    Dim lastrow As Long
    lastrow = ws.Cells(ws.Rows.Count, "A").End(xlUp).row
    ws.Cells(lastrow + 2, "A").Value = "end"
    Set Rng = ws.Range(ws.Range("J1"), ws.Range("J" & Rows.Count).End(xlUp))
    Set Dic = CreateObject("Scripting.Dictionary")
    Dic.CompareMode = 1
        For Each Dn In Rng
            If Not Dic.exists(Dn.Value) Then
                Set Dic(Dn.Value) = CreateObject("Scripting.Dictionary")
            End If
            If Not Dic(Dn.Value).exists(Dn.Offset(, 9).Value) Then
                    Dic(Dn.Value).Add (Dn.Offset(, 9).Value), Dn.Offset(, 19)
                Else
                    Dic(Dn.Value).Item(Dn.Offset(, 9).Value) = Dic(Dn.Value).Item(Dn.
                        Offset(, 9).Value) + Dn.Offset(, 19)
            End If
```

```
        Next Dn
        Set out = ThisWorkbook.Sheets("Country-split")
        c = 1
        out.Range("B1") = "Output"
        For Each F In Dic.keys
            out.Cells(c, "A") = F
            For Each p In Dic(F)
                out.Cells(c, "A") = F
                out.Cells(c, "B") = p
                out.Cells(c, "C") = Dic(F).Item(p)
                c = c + 1
            Next p
            If Not F = "Fund" Then
                c = c + 1
            End If
        Next F
    Application.ScreenUpdating = True
End If
MsgBox "Done"
End Sub

Sub MakeSumCountry()
strPromt = "This macro is step 2 of 5 of calculating the Geographical HHI" & vbNewLine
    & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set ws = Worksheets("Country-split")
    Dim lastrow As Long
    Dim row As Long
    ws.Cells(1, "D").Value = "Total Sum Invested"
    ws.Cells(1, "E").Value = "Squared Investment Share"
    ws.Cells(1, "F").Value = "HHI"
    lastrow = ws.Cells(ws.Rows.Count, "A").End(xlUp).row
    If ws.Cells(lastrow, "A") <> "end" Then
        ws.Cells(lastrow + 2, "A").Value = "end"
    End If
    row = 2
    Do Until ws.Cells(row, 1) = "end"
        If ws.Cells(row, 2) = "" Then
                ws.Cells(row - 1, 4).Value = Sum
                row = row + 1
                Sum = 0
            Else
                Sum = ws.Cells(row, 3).Value + Sum
                row = row + 1
        End If
    Loop
    fundrange = ws.Range(ws.Range("A2"), ws.Range("A" & Rows.Count).End(xlUp))
    ws.Cells(1, "D").Value = "end"
    row = ws.Cells(ws.Rows.Count, "D").End(xlUp).row + 3
    Do Until ws.Cells(row, "D") = "end"
        If ws.Cells(row + 1, "D").Value = "" Then
            x = ws.Cells(row, "D")
            row = row - 1
        End If
        Do While ws.Cells(row, "A") <> "" And ws.Cells(row, "D") <> "end"
            ws.Cells(row + 1, "D").Value = x
            row = row - 1
        Loop
```

```
        ws.Cells(row + 1, "D").Value = x
        If ws.Cells(row, "D") <> "end" Then
            row = row - 1
        End If
    Loop
    ws.Cells(1, "D").Value = "Total Sum Invested"
    'Application.ScreenUpdating = True
End If
MsgBox "Done"
End Sub

Sub shareSquared_Country()
strPromt = "This macro is step 3 of 5 of calculating the Geographical HHI" & vbNewLine
    & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set ws = Worksheets("Country-split")
    fundrange = ws.Range(ws.Range("A2"), ws.Range("A" & Rows.Count).End(xlUp))
    row = 2
    Do Until ws.Cells(row, "A") = "end"
        If ws.Cells(row, "D") = "" Or ws.Cells(row, "D") = 0 Then
                row = row + 1
            Else
                ws.Cells(row, "E").Value = (ws.Cells(row, "C") / ws.Cells(row, "D")) ^
                    2
                row = row + 1
        End If
    Loop
End If
MsgBox "Done"
End Sub

Sub HHI_Country()
strPromt = "This macro is step 4 of 5 of calculating the Geographical HHI" & vbNewLine
    & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'Application.ScreenUpdating = False
    'Application.DisplayAlerts = False
    Set ws = Worksheets("Country-split")
    fundrange = ws.Range(ws.Range("A2"), ws.Range("A" & Rows.Count).End(xlUp))
    row = 2
    For Each fund In fundrange
        If ws.Cells(row, "A") = "" And ws.Cells(row + 1, "A") = "" Then
            GoTo finish
        End If
        s = 0
        s = ws.Cells(row, "E") + s
        Do While ws.Cells(row + 1, "A") = ws.Cells(row, "A") And ws.Cells(row, "A") <>
            "end"
            row = row + 1
            s = ws.Cells(row, "E") + s
        Loop
        row = row + 1
        ws.Cells(row - 1, "F").Value = s
        If ws.Cells(row - 1, "F") = 0 Then
            ws.Cells(row - 1, "F").Delete
        End If
    Next fund
```

```
    'Application.ScreenUpdating = True
    'Application.DisplayAlerts = True
finish:
End If
MsgBox "Done"
End Sub

Sub copyHHI_Country()
strPromt = "This macro is step 5 of 5 of calculating the Geographical HHI" & vbNewLine
    & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'Application.ScreenUpdating = False
    Set inn = Worksheets("Country-split")
    Set out = Worksheets("SDC-split")
    Dim row As Long
    Dim rekke As Long
    row = 2
    rekke = 2
    out.Cells(1, "AP").Value = "HHI_Country"
    Do Until out.Cells(row, "A") = "end"
        Do Until inn.Cells(rekke + 1, "A") = ""
            rekke = rekke + 1
        Loop
        Do While out.Cells(row, "J") = inn.Cells(rekke, "A") And out.Cells(row, "A") <>
             "end"
            out.Cells(row, "AP").Value = inn.Cells(rekke, "F")
            row = row + 1
        Loop
        rekke = rekke + 1
    Loop
End If
'Application.ScreenUpdating = True
MsgBox "Done"
End Sub

Sub StataExport()
strPromt = "This macro creates a new worksheet which will contain all the information
    that we wish to investigate later in Stata" & vbNewLine & vbNewLine & "Run this
    macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Application.DisplayAlerts = False
    On Error Resume Next
    ThisWorkbook.Sheets("StataExport").Delete
    On Error GoTo 0
    Application.DisplayAlerts = True
    Dim ws As Worksheet
    Set ws = Sheets.Add
    ws.Name = "StataExport"
    Set inn = Worksheets("SDC-split")
    Set out = Worksheets("StataExport")
    inn.Select
    inn.Range("A1, B1, F1, J1, O1, AM1").Select
    Selection.Copy
    out.Select
    out.Cells(1, "A").Select
    ActiveSheet.Paste
    out.Cells(1, "C").Value = "Firm Continent"
```

```
out.Cells(1, "G").Value = "Liquidated or Active"
out.Cells(1, "H").Value = "HHI-Industry"
out.Cells(1, "I").Value = "HHI-country"
out.Cells(1, "J").Value = "HHI-Naive"
row = 2
outrow = 2
Do Until inn.Cells(row, "A") = ""
    Do Until inn.Cells(row, "J") <> inn.Cells(row + 1, "J")
        row = row + 1
    Loop
    inn.Select
    inn.Range(inn.Cells(row, "A"), inn.Cells(row, "B")).Select
    Selection.Copy
    out.Select
    out.Cells(outrow, "A").Select
    ActiveSheet.Paste
    out.Cells(outrow, "D").Value = inn.Cells(row, "J")
    out.Cells(outrow, "E").Value = inn.Cells(row, "O")
    out.Cells(outrow, "F").Value = inn.Cells(row, "AK")
    row = row + 1
    outrow = outrow + 1
Loop
row = 2
Do Until out.Cells(row, "A") = ""
    If out.Cells(row, "B") = "United States" Or out.Cells(row, "B") = "Canada" Or
        out.Cells(row, "B") = "Cayman Islands" Or out.Cells(row, "B") = "Bermuda"
        Then
            out.Cells(row, "C").Value = "North America"
        ElseIf out.Cells(row, "B") = "Argentina" Or out.Cells(row, "B") = "Brazil"
            Then
            out.Cells(row, "C").Value = "South America"
        ElseIf out.Cells(row, "B") = "Australia" Or out.Cells(row, "B") = "New
            Zealand" Then
            out.Cells(row, "C").Value = "Oceania"
        ElseIf out.Cells(row, "B") = "Austria" Or out.Cells(row, "B") = "Belgium"
            Or out.Cells(row, "B") = "Channel Islands" Or out.Cells(row, "B") = "
            Czech Republic" Or out.Cells(row, "B") = "Denmark" Or out.Cells(row, "B
            ") = "Finland" Or out.Cells(row, "B") = "France" Or out.Cells(row, "B")
             = "Germany" Or out.Cells(row, "B") = "Greece" Or out.Cells(row, "B") =
             "Ireland" Or out.Cells(row, "B") = "Italy" Or out.Cells(row, "B") = "
            Netherlands" Or out.Cells(row, "B") = "Norway" Or out.Cells(row, "B") =
             "Poland" Or out.Cells(row, "B") = "Portugal" Or out.Cells(row, "B") =
            "Russia" Or out.Cells(row, "B") = "Spain" Or out.Cells(row, "B") = "
            Sweden" Or out.Cells(row, "B") = "Switzerland" Or out.Cells(row, "B") =
             "United Kingdom" Then
            out.Cells(row, "C").Value = "Europe"
        ElseIf out.Cells(row, "B") = "China" Or out.Cells(row, "B") = "Hong Kong"
            Or out.Cells(row, "B") = "India" Or out.Cells(row, "B") = "Indonesia"
            Or out.Cells(row, "B") = "Israel" Or out.Cells(row, "B") = "Japan" Or
            out.Cells(row, "B") = "Malaysia" Or out.Cells(row, "B") = "Philippines"
             Or out.Cells(row, "B") = "Singapore" Or out.Cells(row, "B") = "South
            Korea" Or out.Cells(row, "B") = "Taiwan" Or out.Cells(row, "B") = "
            Thailand" Or out.Cells(row, "B") = "Vietnam" Then
            out.Cells(row, "C").Value = "Asia"
        ElseIf out.Cells(row, "B") = "Egypt" Or out.Cells(row, "B") = "South Africa
            " Then
            out.Cells(row, "C").Value = "Africa"
    End If
    row = row + 1
```

```
    Loop
    row = 2
    Do Until out.Cells(row, "A") = ""
        If out.Cells(row, "F") = "Liquidated" Then
                out.Cells(row, "G").Value = "Liquidated"
            ElseIf out.Cells(row, "F") = "Withdrew" Then
                out.Cells(row, "G").Value = "Withdrew"
            Else: out.Cells(row, "G").Value = "Active"
        End If
        row = row + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub HHI_industry_country_naive()
strPromt = "This macro imports the HHI figures for the industrial, geographical and
    naive HHI into the newly created spreadsheet" & vbNewLine & vbNewLine & "Run this
    macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set inn = Worksheets("SDC-split")
    Set out = Worksheets("StataExport")
    Dim row As Long
    Dim outrow As Long
    row = 2
    outrow = 2
    Do Until inn.Cells(row, "A") = ""
        Do Until inn.Cells(row, "J") <> inn.Cells(row + 1, "J")
            row = row + 1
        Loop
        out.Cells(outrow, "H").Value = inn.Cells(row, "AO")
        out.Cells(outrow, "I").Value = inn.Cells(row, "AP")
        out.Cells(outrow, "J").Value = inn.Cells(row, "AS")
        row = row + 1
        outrow = outrow + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub HHIProduct()
strPromt = "This macro creates a product of all HHI values for each fund" & vbNewLine &
     vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set ws = Worksheets("StataExport")
    row = 2
    ws.Cells(1, "K").Value = "HHIProduct"
    Do Until ws.Cells(row, "A") = ""
        ws.Cells(row, "K").Value = ws.Cells(row, "H") * ws.Cells(row, "I") * ws.Cells(
            row, "J")
        row = row + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub count_IPO()
```

```
strPromt = "This macro Counts number of IPO exits in each fund and divides it by the
    number of portfolio companies to calculate the IPOshare" & vbNewLine & vbNewLine &
    "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'Counts number of IPO exits in each fund and divides it by the number of portfolio
        companies to calculate the IPOshare
    Set inn = Worksheets("SDC-split")
    Set out = Worksheets("StataExport")
    out.Range("L1").Value = "IPO"
    out.Range("M1").Value = "IPOshare"
    out.Range("AC1").Value = "Number of Portfolio Companies"
    Dim IPO As Long
    row = 2
    outrow = 2
    Do Until inn.Cells(row, "A") = ""
        IPO = 0
        compcount = 0
        Do While out.Cells(outrow, "D") = inn.Cells(row, "J")
            If inn.Cells(row, "W") <> "" Then
                compcount = compcount + 1
            End If
            If inn.Cells(row, "W") = "Went Public" Then
                IPO = IPO + 1
            End If
            row = row + 1
        Loop
        out.Cells(outrow, "L").Value = IPO
        out.Cells(outrow, "M").Value = IPO / compcount
        out.Cells(outrow, "AC").Value = compcount
    outrow = outrow + 1
    Loop
End If
MsgBox "Done"
End Sub
Sub count_MA()
strPromt = "This macro counts number of Merger and Acquisition exits in each fund and
    divides it by the number of portfolio companies to calculate the MAshare" &
    vbNewLine & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'Counts number of Merger and Acquistion exits in each fund and divides it by the
        number of portfolio companies to calculate the MAshare
    Set inn = Worksheets("SDC-split")
    Set out = Worksheets("StataExport")
    out.Range("N1").Value = "MA"
    out.Range("O1").Value = "MAshare"
    Dim MA As Long
    row = 2
    outrow = 2
    Do Until inn.Cells(row, "A") = ""
        MA = 0
        compcount = 0
        Do While out.Cells(outrow, "D") = inn.Cells(row, "J")
            If inn.Cells(row, "W") <> "" Then
                compcount = compcount + 1
            End If
            If inn.Cells(row, "W") = "Acquisition" Or inn.Cells(row, "W") = "Merger"
                Then
```

```
                MA = MA + 1
            End If
            row = row + 1
        Loop
        out.Cells(outrow, "N").Value = MA
        out.Cells(outrow, "O").Value = MA / compcount
        outrow = outrow + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub count_LBO()
strPromt = "This macro counts number of Leveraged buyout exits in each fund and divides
     it by the number of portfolio companies to calculate the LBOshare" & vbNewLine &
    vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'Counts number of Leveraged Buyout exits in each fund and divides it by the number
        of portfolio companies to calculate the LBOshare
    Set inn = Worksheets("SDC-split")
    Set out = Worksheets("StataExport")
    out.Range("P1").Value = "LBO"
    out.Range("Q1").Value = "LBOshare"
    Dim LBO As Long
    row = 2
    outrow = 2
    Do Until inn.Cells(row, "A") = ""
        LBO = 0
        compcount = 0
        Do While out.Cells(outrow, "D") = inn.Cells(row, "J")
            If inn.Cells(row, "W") <> "" Then
                compcount = compcount + 1
            End If
            If inn.Cells(row, "W") = "LBO" Then
                LBO = LBO + 1
            End If
            row = row + 1
        Loop
        out.Cells(outrow, "P").Value = LBO
        out.Cells(outrow, "Q").Value = LBO / compcount
        outrow = outrow + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub count_Bankruptcy()
strPromt = "This macro counts number of bankruptcies and defunkt companies in each fund
     and divides it by the number of portfolio companies to calculate the
    Bankruptcyshare" & vbNewLine & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'Counts number of Backruptcies and defunct companies in each fund and divides it by
         the number of portfolio companies to calculate the Bankruptcyshare
    Set inn = Worksheets("SDC-split")
    Set out = Worksheets("StataExport")
    out.Range("AD1").Value = "Bankruptcy"
    out.Range("AE1").Value = "Bankruptcyshare"
```

```
    Dim B As Long
    row = 2
    outrow = 2
    Do Until inn.Cells(row, "A") = ""
        B = 0
        compcount = 0
        Do While out.Cells(outrow, "D") = inn.Cells(row, "J")
            If inn.Cells(row, "W") <> "" Then
                compcount = compcount + 1
            End If
            If inn.Cells(row, "W") = "Bankruptcy - Chapter 7" Or inn.Cells(row, "W") =
                "Bankruptcy - Chapter 11" Or inn.Cells(row, "W") = "Defunct" Then
                B = B + 1
            End If
            row = row + 1
        Loop
        out.Cells(outrow, "AD").Value = B
        out.Cells(outrow, "AE").Value = B / compcount
        outrow = outrow + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub seqnr_and_fundtype()
strPromt = "This macro imports sequence number and fundtype into the worksheet that
    will be used for Stata" & vbNewLine & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'copys information to a separate worksheet that we import into stata.
    Set inn = Worksheets("SDC-split")
    Set out = Worksheets("StataExport")
    Dim row As Long
    row = 2
    outrow = 2
    out.Range("R1").Value = "Seq.Num"
    Do Until inn.Cells(row, "A") = ""
        Do Until inn.Cells(row, "J") <> inn.Cells(row + 1, "J")
            row = row + 1
        Loop
        out.Cells(outrow, "R").Value = inn.Cells(row, "L")
        outrow = outrow + 1
        row = row + 1
    Loop
    row = 2
    outrow = 2
    out.Range("S1").Value = "Fund investment type"
    Do Until inn.Cells(row, "A") = ""
        Do Until inn.Cells(row, "J") <> inn.Cells(row + 1, "J")
            row = row + 1
        Loop
        out.Cells(outrow, "S").Value = inn.Cells(row, "AJ")
        outrow = outrow + 1
        row = row + 1
    Loop
    row = 2
    outrow = 2
    out.Range("T1").Value = "Fund Vintage Year"
    Do Until inn.Cells(row, "A") = ""
```

```
            Do Until inn.Cells(row, "J") <> inn.Cells(row + 1, "J")
                row = row + 1
            Loop
            out.Cells(outrow, "T").Value = inn.Cells(row, "AL")
            outrow = outrow + 1
            row = row + 1
        Loop
    End If
    MsgBox "Done"
    End Sub


    Sub IPO_each_year()
    Set inn = Worksheets("IPOs")
    Set out = Worksheets("StataExport")
    Dim row As Long
    Dim rekke As Long
    out.Range("U1").Value = "Recorded IPOs in that year"
    out.Range("V1").Value = "Change in IPOs"
    row = 2
    rekke = 2
    Do Until inn.Cells(rekke, "A") = ""
        Do Until out.Cells(row, "A") = ""
            If out.Cells(row, "T") = inn.Cells(rekke, "A") Then
                out.Cells(row, "U").Value = inn.Cells(rekke, "B")
                out.Cells(row, "V").Value = inn.Cells(rekke, "C")
            End If
            row = row + 1
        Loop
        row = 2
        rekke = rekke + 1
    Loop
    End Sub

    Sub MSCI()

    strPromt = "This macro imports the MSCI global index as well as 10year geometric
        forward average for each year from 1970 to 2005 " & vbNewLine & vbNewLine & "Run
        this macro?"
    iRet = MsgBox(strPromt, vbYesNo)
    If iRet = vbYes Then
        Set inn = Worksheets("MSCI")
        Set out = Worksheets("StataExport")
        Dim row As Long
        Dim rekke As Long
        out.Range("W1").Value = "MSCI-index"
        out.Range("X1").Value = "Simple return MSCI"
        out.Range("Y1").Value = "Ln-return MSCI"
        out.Range("Z1").Value = "MSCI Ten Year Fwd Avg."
        row = 2
        rekke = 8
        Do Until inn.Cells(rekke, "A") = ""
            Do Until out.Cells(row, "A") = ""
                If out.Cells(row, "T") = inn.Cells(rekke, "A") Then
                    out.Cells(row, "W").Value = inn.Cells(rekke, "B")
                    out.Cells(row, "X").Value = inn.Cells(rekke, "C")
                    out.Cells(row, "Y").Value = inn.Cells(rekke, "D")
                    out.Cells(row, "Z").Value = inn.Cells(rekke, "E")
                End If
```

```
            row = row + 1
        Loop
        row = 2
        rekke = rekke + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub GDP()
strPromt = "This macro imports the global GDP growth as well as 10year geometric
    forward average for each year from 1961 to 2003 " & vbNewLine & vbNewLine & "Run
    this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set inn = Worksheets("GDP")
    Set out = Worksheets("StataExport")
    Dim row As Long
    Dim rekke As Long
    out.Range("AA1").Value = "GDP growth"
    out.Range("AF1").Value = "GDP 10yr GeoAvg."
    row = 2
    rekke = 5
    Do Until inn.Cells(rekke, "A") = ""
        Do Until out.Cells(row, "A") = ""
            If out.Cells(row, "T") = inn.Cells(rekke, "A") Then
                out.Cells(row, "AA").Value = inn.Cells(rekke, "K")
                out.Cells(row, "AF").Value = inn.Cells(rekke, "J")
            End If
            row = row + 1
        Loop
        row = 2
        rekke = rekke + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub firmxp()
strPromt = "This macro calculates the years of experience each general partner firm had
     at the begining of each fund's vintage year." & vbNewLine & vbNewLine & "Run this
    macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    'calculates the years of experience each general partner firm had at the begining
        of each fund's vintage year.
    Set ws = Worksheets("StataExport")
    Dim row As Long
    Dim rekke As Long
    Columns("A:A").Select
        ActiveWorkbook.Worksheets("StataExport").Sort.SortFields.Clear
        ActiveWorkbook.Worksheets("StataExport").Sort.SortFields.Add Key:=Range( _
            "A2:A3903"), SortOn:=xlSortOnValues, Order:=xlAscending, DataOption:= _
            xlSortNormal
        With ActiveWorkbook.Worksheets("StataExport").Sort
            .SetRange Range("A1:AG3903")
            .Header = xlYes
            .MatchCase = False
            .Orientation = xlTopToBottom
```

```
            .SortMethod = xlPinYin
            .Apply
        End With
    ws.Range("AB1").Value = "Firm Experience"
    row = 2
    rekke = row
    Do Until ws.Cells(row, "A") = ""
        Start = ws.Cells(row, "T")
        If ws.Cells(row, "A") <> ws.Cells(row - 1, "A") And ws.Cells(row, "A") <> ws.
            Cells(row + 1, "A") Then
            experience = ws.Cells(row, "T").Value - Start
            ws.Cells(row, "AB") = experience
            row = row + 1
            GoTo skip
        End If
        If ws.Cells(row, "A") = ws.Cells(row + 1, "A") Or ws.Cells(row, "A") = ws.Cells
            (row - 1, "A") Then
            rekke = row
            Start = ws.Cells(rekke, "T")
        End If
        Do While ws.Cells(rekke, "A") = ws.Cells(rekke - 1, "A") Or ws.Cells(rekke, "A"
            ) = ws.Cells(rekke + 1, "A")
            If ws.Cells(rekke, "T") < Start Then
                Start = ws.Cells(rekke, "T")
            End If
            If ws.Cells(rekke, "A") <> ws.Cells(rekke + 1, "A") Then
                GoTo skiprekke
            End If
            rekke = rekke + 1
        Loop
skiprekke:
        Do While ws.Cells(row, "A") = ws.Cells(row - 1, "A") Or ws.Cells(row, "A") = ws
            .Cells(row + 1, "A")
            experience = ws.Cells(row, "T") - Start
            ws.Cells(row, "AB").Value = experience
            row = row + 1
            If ws.Cells(row, "A") <> ws.Cells(row - 1, "A") Then
                GoTo skip
            End If
        Loop
        row = row + 1
skip:
    Loop
End If
MsgBox "Done"
End Sub

Sub yield()
strPromt = "This macro imports the interest rates for 10 year US Treasury bonds and 1
    year T-bills." & vbNewLine & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set inn = Worksheets("Spread")
    Set out = Worksheets("StataExport")
    Dim row As Long
    Dim rekke As Long
    out.Range("AG1").Value = "Bond yield ten year"
    out.Range("AH1").Value = "T-bill yield one year"
    out.Range("AI1").Value = "Spread Bond - Bill"
```

```
    row = 8
    rekke = 2
    Do Until inn.Cells(row, "A") = ""
        rekke = 2
        Do Until out.Cells(rekke, "T") = ""
            If inn.Cells(row, "A") = out.Cells(rekke, "T") Then
                out.Cells(rekke, "AG").Value = inn.Cells(row, "F")
                out.Cells(rekke, "AH").Value = inn.Cells(row, "C")
                out.Cells(rekke, "AI").Value = inn.Cells(row, "H")
            End If
            rekke = rekke + 1
        Loop
    row = row + 1
    Loop
End If
MsgBox "Done"
End Sub

Sub InvestedCompaniesbyFirm()
'count number of portfolio companies in each fund
Set ws = Worksheets("NaiveHHI")
ws.Range("G1").Value = "Number of portfolio companies invested by fund"
row = 3
Do Until ws.Cells(row, "A") = "" And ws.Cells(row + 1, "A") = ""
    num = 0
    Do Until ws.Cells(row, "A") = ""
        If ws.Cells(row, "A") = ws.Cells(row + 1, "A") Or ws.Cells(row + 1, "A") = ""
            Then
            num = num + 1
        End If
        row = row + 1
    Loop
    ws.Cells(row - 1, "G").Value = num
    row = row + 1
Loop
'Links number of portfolio companies to each fund in "StataExport" sheet.
End Sub

Sub copynum()
Set out = Worksheets("StataExport")
Set inn = Worksheets("NaiveHHI")
Dim row As Long
Dim rekke As Long
row = 2
rekke = 2
Do Until inn.Cells(row, "A") = "" And inn.Cells(row + 1, "A") = ""
    rekke = 2
    Do Until inn.Cells(row + 1, "A") = ""
        row = row + 1
    Loop
    Do Until out.Cells(rekke, "A") = ""
        If inn.Cells(row, "A") = out.Cells(rekke, "D") Then
            out.Cells(rekke, "V").Value = inn.Cells(row, "G")
            GoTo nextfund
        End If
        rekke = rekke + 1
    Loop
nextfund:
    row = row + 1
```

```
Loop
End Sub


Sub NumCompInvestedSoFar()
Set ws = Worksheets("StataExport")
Dim row As Long
ws.Range("U1").Value = "Invested portfolio companies so far"
Columns("A:AI").Select
    ActiveWorkbook.Worksheets("StataExport").Sort.SortFields.Clear
    ActiveWorkbook.Worksheets("StataExport").Sort.SortFields.Add Key:=Range( _
        "A2:A3903"), SortOn:=xlSortOnValues, Order:=xlAscending, DataOption:= _
        xlSortNormal
    ActiveWorkbook.Worksheets("StataExport").Sort.SortFields.Add Key:=Range( _
        "T2:T3903"), SortOn:=xlSortOnValues, Order:=xlAscending, DataOption:= _
        xlSortNormal
    With ActiveWorkbook.Worksheets("StataExport").Sort
        .SetRange Range("A1:AI3903")
        .Header = xlYes
        .MatchCase = False
        .Orientation = xlTopToBottom
        .SortMethod = xlPinYin
        .Apply
    End With
row = 2
Do Until ws.Cells(row, "A") = ""
    Start = 0
    ws.Cells(row, "U").Value = Start
    Do Until ws.Cells(row, "A") <> ws.Cells(row + 1, "A")
        ws.Cells(row, "U").Value = Start
        Start = Start + ws.Cells(row, "V")
        row = row + 1
        If ws.Cells(row, "A") <> ws.Cells(row + 1, "A") Then
            ws.Cells(row, "U").Value = Start
        End If
    Loop
    row = row + 1
Loop
End Sub

Sub PE()
strPromt = "This macro imports the price/earnings and the PE10 for the S&P500." &
    vbNewLine & vbNewLine & "Run this macro?"
iRet = MsgBox(strPromt, vbYesNo)
If iRet = vbYes Then
    Set ws = Worksheets("PriceEarnings")
    Set out = Worksheets("StataExport")
    out.Range("AJ1").Value = "PE10"
    row = 9
    rekke = 9
    Do Until ws.Cells(row, "A") = ""
        If ws.Cells(row, "B") = 1 And ws.Cells(row - 1, "B") = 9 Then
            ws.Cells(row, "B").Value = 10
        End If
        row = row + 1
    Loop
    row = 9
    Do Until ws.Cells(row, "A") = ""
        If ws.Cells(row, "B") <> 1 Then
```

```
            ws.Rows(row).Delete
            GoTo skip
        End If
        row = row + 1
skip:
    Loop
    row = 9
    Do Until ws.Cells(row, "A") = ""
    rekke = 2
        Do Until out.Cells(rekke, "T") = ""
            If ws.Cells(row, "A") = out.Cells(rekke, "T") Then
                'The PE10 is based on 10 previous years, so we shift it by 10 years
                    forward, to get the Price/Earnings average of the typical lifecycle
                     of a PE-fund
                out.Cells(rekke, "AJ").Value = ws.Cells(row + 10, "L")
            End If
            rekke = rekke + 1
         Loop
    row = row + 1
    Loop
    row = 9
    out.Range("AK1").Value = "PE"
    Do Until ws.Cells(row, "A") = ""
        rekke = 2
        Do Until out.Cells(rekke, "T") = ""
            If ws.Cells(row, "A") = out.Cells(rekke, "T") Then
                out.Cells(rekke, "AK").Value = ws.Cells(row, "M")
            End If
            rekke = rekke + 1
         Loop
    row = row + 1
    Loop
MsgBox "Done"
End Sub
```

8 Bibliography

Baum, C. F. Stata Tip 63: Modeling Proportions. *Stata Journal*, 8(2):299–303, 2008.

Bienz, C. and Leite, T. E. A Pecking Order of Venture Capital Exits. *Unpublished Working Paper*, pages 1–17, 2008.

Bottazzi, L., Da-Rin, M., and Hellmann, T. The Changing Face of the European Venture Capital Industry: Facts and Analysis. *Journal of Private Equity*, 2004.

Brigl, M., Herrera, A., Liechtenstein, H., Meerkatt, H., Prats, J., and Rose, J. The Advantage of Persistence - How the Best Private-Equity Firms "Beat the Fade". Boston, MA, 2008.

Cameron, C. A., Gelbach, J. B., and Miller, D. L. Robust Interference with Multi-Way Clustering. Technical report, National Bureau of Economic Research, Cambridge, MA, 2006.

Cressy, R., Munari, F., and Malipiero, A. Playing to Their Strengths? Evidence that Specialization in the Private Equity Industry Confers Competitive Advantage. *Journal of Corporate Finance*, 13(4):647–669, 2007.

Diller, C. and Kaserer, C. What Drives Cash Flow Based European Private Equity Fund Returns?–Fund Inflows, Skilled GPs and/or Risk? *LSE/RICAFE Working Paper*, 15, 2005.

Doidge, C., Karolyi, G. A., and Stulz, R. M. The US Left Behind: The Rise of IPO Activity Around the World. Technical report, National Bureau of Economic Research, 2011.

EVCA. 2013 European Private Equity Activity. Technical report, EVCA, 2014.

Ewens, M., Jones, C. M., and Rhodes-Kropf, M. The Price of Diversifiable Risk in Venture Capital and Private Equity. *Review of Financial Studies*, 26(8):1854–1889, 2013.

FED. 2014. Bord of Governors of the Federal Reserve System - Downloaded from http://www.federalreserve.gov/releases/h15/data.htm (28.03.2014).

Gilligan, J. and Wright, M. *Private Equity Demystified*. ICA Corporate Finance Faculty, 2010.

Gompers, P. and Lerner, J. *The Venture Capital Cycle*. The MIT Press, Cambridge MA, 2004.

Gompers, P., Kovner, A., Lerner, J., and Scharfstein, D. Venture Capital Investment Cycles: The Impact of Public Markets. *NBER Working Paper Series*, 2005.

Gompers, P. A. Optimal Investment, Monitoring, and the Staging of Venture Capital. *The Journal of Finance*, 1995.

Gottschalg, O., Phalippou, L., and Zollo, M. Performance of Private Equity Funds: Another Puzzle? Working paper series, 2004.

Jensen, M. C. Eclipse of the Public Corporation. *Harvard Business Review*, 67(5):61–74, 1989.

Judge, G. G., Griffiths, W. E., Hill, C., Lütkepohl, H., and Lee, T.-C. *The Theory and Practice of Econometrics*. Wiley, New York, 2 edition, 1985.

Kaplan, S. The Effects of Management Buyouts on Operating Performance and Value. *Journal of Financial Economics*, 1989.

Kaplan, S. N. and Schoar, A. Private Equity Performance: Returns, Persistence, and Capital Flows. *The Journal of Finance*, 2005.

Kenyon, H. Prequin Quarterly Private Equity - Q3. Technical report, Prequin Ltd, October 2012.

Kieschnick, R. and McCullough, B. D. Regression Analysis of Variates Observed on (0, 1): Percentages, Proportions and Fractions. *Statistical modelling*, 3(3):193–213, 2003.

Ljungqvist, A. and Richards, M. The Cash Flow, Return and Risk Characteristics of Private Equity. Unpublished work, 2003.

Ljungqvist, A., Richardson, M., and Wolfenzon, D. The Investment Behaviour of Buyout Funds: Theory and Evidence. Working paper, 2007.

Lossen, U. The Performance of Private Equity Funds: Does Diversification Matter? Discussion paper series of sfb/tr 15 governance and the efficiency of economic systems, University of Munich, 2006.

Müller, K. *Investing in Private Equity Partnerships*. Betriebswirtschaftlicher Verlag, Wiesbaden, 1 edition, 2008.

Norton, E. and Tenenbaum, B. Specialization Versus Divercification as a Venture Capital Investment Strategy. *Journal of Business Venturing*, 1993.

Papke, L. and Wooldridge, J. Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates. *Journal of applied econometrics*, 11:619–632, 1996.

Petersen, M. A. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. Northwestern University, 2008.

Ramalho, E. A., Ramalho, J., and Murteira, J. Alternative Estimating and Testing Empirical Strategies for Fractional Regression Models. *Journal of Economic Surveys*, 25(1):19–68, 2011.

Ramalho, J. and Silva, J. V. A Two-Part Fractional Regression Model for the Financial Leverage Decisions of Micro, Small, Medium and Large Firms. *Quantitative Finance*, 9(5):621–636, 2009.

Ritter, J. R. and Welch, I. A Review of IPO Activity, Pricing, and Allocations. *The Journal of Finance*, 2002.

Roberts, M. R. and Whited, T. M. Endogeneity in Emperical Corporate Finance. Unpublished Work, 2012.

Schiller, R. J. Irrational Exuberance. *Princeton UP*, 2000. Downloaded from http://www.econ.yale.edu/~shiller/data/ie_data.xls (14.05.2014).

Stock, J. H. and Watson, M. M. *Introduction to Econometrics*. Pearson Education, Essex, 2011.

Thompson, S. Simple Formulas for Standard Errors that Cluster by Both Firm and Time. *Journal of Financial Economics*, 2010.

Wooldridge, J. M. *Introductory Econometrics: A Modern Approach*. South Western, Mason, OH, 5 edition, 2012.