Norwegian School of Economics
Bergen, Spring 2015

Understanding Private Equity Performance

A Review of Dynamics Driving Fund Performance

Jon Olav Aurdal & Ørjan Mydland Grønsund

Supervisor: Tyler Hull

Master of Science in Economics and Business Administration,
Finance

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

In recent years, more papers on private equity performance have emerged, casting light over a market that earlier was characterised by privacy and secrecy. Early studies mostly use VentureXpert as a data provider, however, data from this provider has been under a lot of criticism lately, and new data providers have emerged. In this thesis, we study the performance of buyout and venture funds from 1990 to 2008 using a dataset from Preqin. Previous studies have mostly focused on IRR or a modification of this metric. We have compared these findings with our results and use a widely reported investment multiple to see if there are discrepancies that can explain the differences in results. Based on findings from other papers, the dataset is of high quality and is less prone to bias compared to datasets previously used in private equity research.

In our study of fund types, we see a general tendency of buyout outperforming venture. We have also looked more closely at sequence numbers and see that there is a negative correlation between performance and sequence numbers. This suggests that experience is not necessarily a contributing factor for good performance. We find indications that past performance may be well suited for risk reduction, but is not necessarily indicative for future performance.

As Preqin has been little used in private equity research, our results contribute to this field by showing that Preqin, as a data provider, is well suited for academic research. We also test the validity of past research and show that, although the concepts are still valid, an update based on newer data points is warranted.

Preface

This thesis is written as a part of the master's degree in finance at the Norwegian School of Economics (NHH).

Our choice of private equity as a topic was mainly due to the fascination of this fairly secretive asset class. Several interesting guest lectures, emphasising the complex nature of private equity, really sparked our interest in this field. We discovered that much of the research on private equity were focused on events happening before and around the dot-com bubble. As the 2000s saw a lot of important economic changes, we felt that there was time for a review of previous established consensuses.

Working on this paper has been an interesting an engaging journey. It has been both frustrating and challenging at times, but on overall, an enriching and extremely valuable experience.

We would like to thank our supervisor Tyler Hull for his help in completing this thesis and pointing us in the right direction early on. We would also like to thank Carsten Bienz for giving us access to the data. Justin Kimble, an account manager at Preqin and our direct contact at the firm, also deserves thanks for helping us in navigating the database and answering our, sometimes, stupid questions. Last but not least, we thank our families and friends for all support they have given us these last couple of months.

Bergen, 2015

Jon Olav Aurdal

Ørjan Mydland Grønsund

Contents

List of Figures and Tables

Figures

Tables

List of Acronyms

CA	Cambridge Associates
DPI	Distributions to Paid-In capital
GIPS	Global Investment Performance Standards
GP	General Partner
IRR	Internal Rate Of Return
LPs	Limited Partner
MWW	Mann-Whitney Wilcoxon
NAV	Net Asset Values
PE	Private Equity
PIC	Paid-in capital to Committed capital
RVPI	Residual Value to Paid-in capital
t-test	Independent Student's t-test
TVPI	Total Value to Paid-In capital
VE	VentureXpert/Thomson Reuters/Thomson Venture Economics

1 Introduction

Despite the relative size and recent booms in private equity it is still, despite academic and practitioner research, an asset class shrouded in mystery. By definition, private equity is private, and the asset class has been able to keep much information hidden from researchers, colleagues, rivals, authorities and the general public.

The recent changes in legislation has made data collection easier to obtain and there are now better, more reliable data sources, than what has historically been the case. Our data is provided by Preqin and it displays potential benefits over previously used data. Easier data access in combination with the recent rise in popularity of the asset class, has resulted in numerous papers trying to ascertain the risks and rewards associated with private equity investments.

Our study focuses mainly on the constituent factors that drive the performance of private equity. The study is limited to buyout and venture, and each of them will be analysed both separately and in combination.

Our thesis has the following structure. In section 2, we present the basics of private equity followed by a review of past literature on the field. In section 4, we cover some of the main theory on private equity performance, and in section 5, we look more closely at the data on hand. In the next sections we first present our hypotheses and then explain the methodology used. In section 8, we discuss our findings before summarising our results in section 9.

2 Background

2.1 Private Equity

In theory, the private equity (PE) term refers to a market opposite of the more known public market. PE in general covers investment strategies like venture capital, mezzanine, buyout and real estate to mention some. In later years, the term is used to refer to *"later-stage development capital, but mostly buy-outs and buy-ins of established businesses"*(Gilligan & Wright, 2014, p. 14). We will use the more general interpretation of this term in our thesis, a term covering both early and later stage investments.

Investing in PE is mostly done through PE funds. These funds are run by fund managers, also called *general partners* (GPs), while the funds' investors are called *limited partners* (LPs). Once a PE fund is created, it starts seeking investors, entering a period called *on the road*. In this period, investors commit money to the fund, entering in to an LP agreement. When enough money is raised, the fund is officially *closed*, and the GPs can start investing. This is true in most cases, but there are instances where closed funds have been reopened.

PE funds have limited lifespans. The first five to six years are most often used to invest, hence this period is called an *investment period*. After investments are made, GPs focus on getting the best results possible, often trough strategies like restructuring and active ownership, before exiting them. In this last period, no new investments are made, only follow-up investments in their portfolio companies. Usually the pre-agreed length of a fund is approximately 10 years, with a two-year possible extension. A PE company will in most cases always have a fund in the investment period (Gilligan & Wright, 2014). Hence new funds are, on average, created every three years (Kaplan & Schoar, 2005).

The full amount of capital LPs commit to a fund over its lifespan is called *committed capital*. The committed capital is given to a fund on either a fixed schedule or when a fund calls for it. This is called a *capital call*, and the total amount of money available to a fund at any given time is called *dry powder*. If an LP is not able to pay a fund when a capital call comes, it is often arranged so that the other LPs must cover this amount (Gilligan & Wright, 2014). The contract details are different from fund to fund, but all are stipulated in the investment contract.

2.2 Fees

PE investments do not realise immediate returns, but the costs of running a PE fund start right away. Salaries need to be paid and due diligences need to be done. In order to handle these costs, LPs pay management fees to GPs. These fees are annual and approximately 2% of committed capital, and usually management fees are reduced as a fund exits the investment period and starts realising returns.

Even though management fees are much needed, some argue that they may cause a principal-agent problem. As a fund grows, so does fees, giving GPs a larger profit independent of fund performance (Gilligan & Wright, 2014).

There is also a second form of compensation for PE funds called *carried interest*. When a fund's lifetime is over, GPs gain a certain share of the profit after committed capital is paid back. 20% carried interest is most common, and usually accounts for a GP's biggest profit (Metrick & Yasuda, 2011). There are several variations of compensations more complex and detailed than the ones presented here, but it is not in the essence of this thesis to dissect compensation schemes.

2.3 General Partners

GPs may refer to a whole company, but may also just refer to a team of individuals within a company having responsibility for a particular fund. Their first task is to raise money in order to gain capital needed for investing. Once a fund is closed, GPs can start looking for their first investments.

Before investments are made, GPs have to structure financing and negotiate terms in order to close a deal. When a deal is finally closed, an investment has to be closely monitored and actively managed if necessary/possible. In the end, GPs exit investments, realising their returns. We mentioned earlier that there might be some principal-agent problems relating to PE. As a preventive factor, most GPs usually invest in the fund themselves. About one per cent of a fund's capital come from GPs, increasing their incentive to perform well (Metrick & Yasuda, 2011). There is also the carried interest, which is potentially huge for a profitable firm.

2.4 Limited Partners

The reason investors are called LPs, is that they have limited liability and thus cannot lose more money than they invest. A fund is usually set up as a separate limited life partnership, preventing any double taxation, making it more attractive to possible investors. There are a lot of different LPs, but pension funds, both public and private, are by far the biggest of them. Following pension funds, we find foundations, fund of funds companies, insurance companies and endowment plans to mention some (Gilligan & Wright, 2014).

2.5 Fund Types

When capital is called, it is time for GPs to start investing. Different funds are classified according to the investments they make. Buyout, measured in number of funds, size of deals, and size of exits are the biggest fund type in PE (Gilligan & Wright, 2014; Metrick & Yasuda, 2011). Venture is the second largest, and these two fund types are most researched.

Buyout funds often take majority control of the companies they invest in, and usually these companies are well established. In contrast, venture funds take smaller stakes in companies. These companies are often newly started or seen upon as up and coming, making the deals smaller than those of buyout funds. However, venture funds may realise bigger returns on successful exits.

2.6 Returns

The returns and performance of PE funds are measured in many ways. According to Global Investment Performance Standards (GIPS), presentations of some measures are mandatory as of each annual period end (CFA Institute, 2010):

- Paid-In capital
- Distributions
- Committed Capital
- Total Value to Paid-In capital (TVPI)
- Distributions to Paid-In capital (DPI)
- Paid-In capital (PIC)
- Residual Value to Paid-In capital (TVPI)

- Internal Rate of Return (IRR[1])

These latest standards come from 2011, and even though some of these measures are reported for funds decades back, secrecy and privacy have been a problem for PE research. Earlier, it was not demanded that fair value was used in calculations and many funds were reluctant to give up anything but final returns. This secrecy, among other things, has led to an increasing number of research papers on PE in the last 10-15 years.

IRR is perhaps the most popular performance measure. This is the annualised yield of the investments' underlying cash flow. The main advantage of IRR is that it considers timing of cash flows. However, the metric does provide some drawbacks that will be illuminated later.

Other popular performance metrics are investment multiples. PE funds have, even before the GIPS requirements, reported multiples. These can, together with IRR, be used to get a better understanding of the true returns LPs get from their investments. IRR and multiples are complementary and both should be used with caution when reviewing fund performance and in comparisons of PE performance.

[1] All references to IRR is net of fees (Net IRR) unless stated otherwise.

3 Litterature Review

As mentioned, more and more papers on PE have been published in later years. Many of these papers focus on whether PE funds perform better than a public market or not. Phalippou and Gottschalg (2008) find evidence that the S&P 500 outperforms PE, net-of-fees, by 3% per year. Robinson and Sensoy (2011) on the other hand, find evidence of the opposite. Compared to the S&P 500, and seen over a fund's lifetime, buyout funds outperform the index by 18%, while venture funds outperform it by 3%. A combination of these results are found in Kaplan and Schoar (2005), who reports that buyout funds underperform compared to the public market. They also find that venture funds underperform if returns are equally weighted, and overperform if returns are weighted by capital. Both Kaplan and Schoar & Phalippou and Gottschalg find that PE outperforms the public market gross-of-fees.

There are several reasons why results differ, but one important factor seems to be the choice of dataset. The most common data providers are:

- Burgiss
- Cambridge Associates (CA)
- Preqin
- VentureXpert/Thomson Reuters/Thomson Venture Economics (VE)

Earlier papers mostly use VE. Although Preqin and CA have been around for some time, Preqin has not been used a lot in PE research (Harris, Jenkinson, & Stucke, 2010). The most recent data provider is Burgiss, which have gained popularity among researchers in later years.

As mentioned, most of the earlier papers, including both Kaplan and Schoar (2005) and Phalippou & Gottschalg (2008), use VE. Later research have shown that the VE data has several negative features. Stucke (2011) finds that net asset values (NAV) and cash flows were not updated for years. NAVs was rolled on for each year, making the numbers going forward almost meaningless at the end. These NAVs are for instance used by Phalippou & Gottschalg. As a funds maturity increases, IRR will decrease, thus understating returns. This affects Kaplan and Schoar, Phalippou and Gottschalg and most other papers based on VE.

In recent years, several papers have started to evaluate the different datasets, comparing them against each other. Harris, Jenkinson, & Kaplan (2014) show that Preqin, CA and Burgiss have more or less the same performance results. They also find evidence suggesting these datasets are unbiased, and hence suitable for academic research. Further, they find similar results for VE as Stucke (2011). So do Harris, Jenkinson & Stucke (2010). They compare VE, Preqin and CA, but unlike Harris, Jenkinson & Kaplan, find big differences in the datasets. The various mixes of fund types and small coverages of total funds are a big concern and "*Some of these differences are not readily explained by random variation and suggest systematic effects related to data methods and sample selection*" (Harris et al., 2010, p. 24). They also find that Preqin usually have higher performance figures, especially for venture in the early 90s. Other reasons why results differ may be the research period. Performance in PE is very cyclical, and there might be big differences in results (Higson & Stucke, 2012). In addition, different definitions or classifications in datasets may have an impact on results.

Buyout and venture are by far the most researched fund types. In Kaplan & Schoar (2005), venture is generally the better performing fund type. Hsu (2004) finds that venture companies with a high reputation have a better chance of getting their offers accepted than those with lower reputation. In addition, high reputation leads to better deals, increasing chances of higher IRRs. While other studies also find that venture outperform buyout, this is only for smaller periods (Harris, Jenkinson, & Kaplan, 2014). Looking at the whole sample period, Ljungqvist & Richardson (2003), Robinson & Sensoy (2011) and Harris, Jenkinson & Kaplan (2014) all find better performance for buyout compared to venture. The two latter use a similar sample period as we do, making their results comparable to ours.

There is not a lot of research on how funds are performing depending on whether they are raised in a boom or bust period. Kaplan & Schoar (2005) find evidence that funds raised in periods of high economic growth are less likely to create follow-on funds, implying they perform worse than funds raised in bust periods. According to Robinson & Sensoy (2011), low performance in PE follows periods with high fundraising. Barber and Yasuda (2014) find that when interim performance of a fund is peaking, GPs start fundraising. Typically, one would assume this happens towards the middle or end of a boom period, hence fundraising will peak close to the next bust period. This could partly explain the results of both Kaplan & Schoar and Robinson & Sensoy.

Comparisons of decades could be helpful in detecting possible changes that the large inflows of institutional investors and low cost of capital during the 00s, had on the PE industry (Appelbaum & Batt, 2012). As we analyse data up until 2008, we can compare the 90s against the 00s. Since we have funds in our dataset that has not yet been liquidated, results of this comparison may be in favour of the 90s. Ljungqvist & Richardson point out that the average IRR of a fund does not turn positive until year eight. Hence, fund performance for the 00s may be biased downwards. This theory is also supported by Steer and Ellis (2011). In their study of PE valuations, they find that even though interim IRRs can be overstated, they tend to be downwardly biased. However, this downward bias becomes insignificant when a fund reaches year seven, meaning that there is no systematic evidence of bias in valuations once a fund is sufficiently mature.

Harris, (2014) report high performance for buyout throughout both decades, but for venture funds, performance in the 00s was low. Higson & Stucke (2012) only study buyout funds, but find a significant downward trend in performance through vintages. The results of these two papers suggest better performance in the 90s compared to the 00s.

Controlling for fund size, performance vary a lot in different papers. Kaplan & Schoar (2005) and Higson & Stucke (2012) find that larger funds perform better than small funds. The former also finds that past performance is positively related to capital inflows. Hence, funds will generally increase after periods of excessive economic growth, and well-performing GPs will tend to raise larger follow-on funds. These findings are consistent with an early study by Sirri & Tufano (1998) on mutual funds and more recent studies by Kaplan & Strömberg (2005) and Robinson & Sensoy (2011) on PE.

Further, Robinson & Sensoy (2011) reports lower returns after periods of high fundraising. If this effect dominates, we could see smaller funds performing better, resulting in size being negatively correlated with returns. This is supported by Ljungqvist & Richardson (2003) and Gompers & Lerner (2000). Gompers & Lerner’s results indicate that funds pay a higher price for their investments following capital inflows. Thus, chances are that smaller funds will be able to outperform larger funds. This may also be one of the reasons why successful GPs choose not to increase follow-on fund sizes (Kaplan & Schoar, 2005). Hellman & Puri (2002) find that GPs focusing on venture can have a good influence on the outcome of the investments by using their skills and knowledge. Metrick & Yasuda (2010) agree with this relating to venture, and add that buyout is more scalable, implying past performance has different implications for different fund types. Higson & Stucke (2012) suggest that larger buyout funds may perform better because they get easier access to debt financing, often at more favourable terms than smaller funds.

While comparisons of fund size and performance are well researched, fewer look at sequence number and performance. According to Kaplan & Schoar (2005), first-time funds perform worse than later funds, but looking at specific GPs, higher sequence numbers results in lower performance. The latter part probably coincides with the fact that high performing funds are more likely to have follow-on funds (Chung, Sensoy, Stern, & Weisbach, 2012). Harris, Jenkinson, Kaplan, & Stucke (2014) confirm this, but dig even deeper into the differences between fund types and decades. GPs with well performing buyout funds pre-2000, seem to raise new well-performing funds, but persistence is not found post-2000. For GPs with venture funds, persistence is found in both periods. Looking at sequence number and size together, Kaplan & Schoar find evidence suggesting that "*funds with persistently good performance are especially favored in the fund raising process*" (Kaplan & Schoar, 2005, p. 21). This implies that fund size increases with sequence number as raising new funds are most often done by those GPs with already successful funds.

Metrick & Yasuda (2010) have also taken a closer look at fund size and sequence numbers. They find that for buyout funds, GPs with experience increase the fund size sharply even though they know this will result in worse performance. Larger funds result in higher fees, which again results in higher GP income. They actually reduce chances of raising more follow-on funds in favour of short-term income by making their next fund larger. As mentioned earlier, this might be easier with buyout funds as they are more scalable than venture funds. Such behaviour may cause principal-agent problems, where GPs favour higher fees at the expense of LP profitability. They do not find similar returns for venture funds.

4 Theory

4.1 Performance Metrics

4.1.1 Multiples

Depending on the multiple, and how it is calculated, unrealised returns may be included. There are uncertainty regarding these multiples, which is especially true for figures reported by GPs before the introduction of GIPS private equity provisions. These standards have clear definitions of multiple reporting, which make them better for comparisons.

4.1.1.1 Distributions to Paid-In capital

DPI is a realisation multiple that provides additional information as to how much of the return that has actually been realised and distributed to LPs. DPI is given by:

$$DPI = \frac{Cumulative\ distributions}{Pain - in - capital} \tag{1}$$

In a fund's early life, this multiple is typically zero since there has been no realisations yet. However, it will grow over a fund's life. This metric may be very volatile towards the end, as a fund may call on more capital to reinvest in portfolio companies before exit. When a fund's DPI equals one, this is the LPs brake-even point. However, DPI is presented in nominal terms since time value of money is not factored in.

DPI and TVPI are the same after a fund has been liquidated and can be an important multiple in comparing PE firms. It gives a measure of how much is actually realised, and in the end realised returns are what matters.

4.1.1.2 Residual Value to Paid-In capital

GIPS private equity provisions also require the presentation of RVPI. RVPI is a measure on how much of the return is unrealised, and is the counterpoint to DPI. As a fund matures, RVPI will increase to a peak, and eventually decrease to a residual market value of zero when a fund is liquidated. At this point, the entire return of the fund has been distributed.

$$RVPI = \frac{Net\ Asset\ Value}{Paid\ \ In\ \ capital} \tag{2}$$

This metric is subjective as there are multiple ways a fund can calculate the value of unrealised returns. The guidelines provide a broad foundation for valuing assets, and aim to improve comparability between GPs. They recommend a concept of fair value, which is the amount an asset could be sold for or acquired by, in a transaction between willing and unrelated parties. It is an estimate of likely exchange price and does involve subjective judgements. Hence, there is a potential to manipulate these numbers.

However, research on UK venture and PE valuations, suggests that there is little sign of upwards systematic bias in interim valuations of unrealised returns. This might suggest that RVPI is useful when combined with DPI in evaluations of PE performance (Steer & Ellis, 2011).

4.1.1.3 Total Value to Paid-In capital

The standards require funds to report TVPI. This multiple is also known as an investment multiple and is the sum of DPI and RVPI. It is also given by:

$$TVPI = \frac{(Net\ asset\ value + Distributions)}{Paid\ \ In\ \ capital} \tag{3}$$

The metric gives an overall performance of a PE fund and is the most used multiple of return. For relatively young funds, and inexperienced GPs, TVPI might be highly uncertain.

4.1.2 Drawbacks of multiples

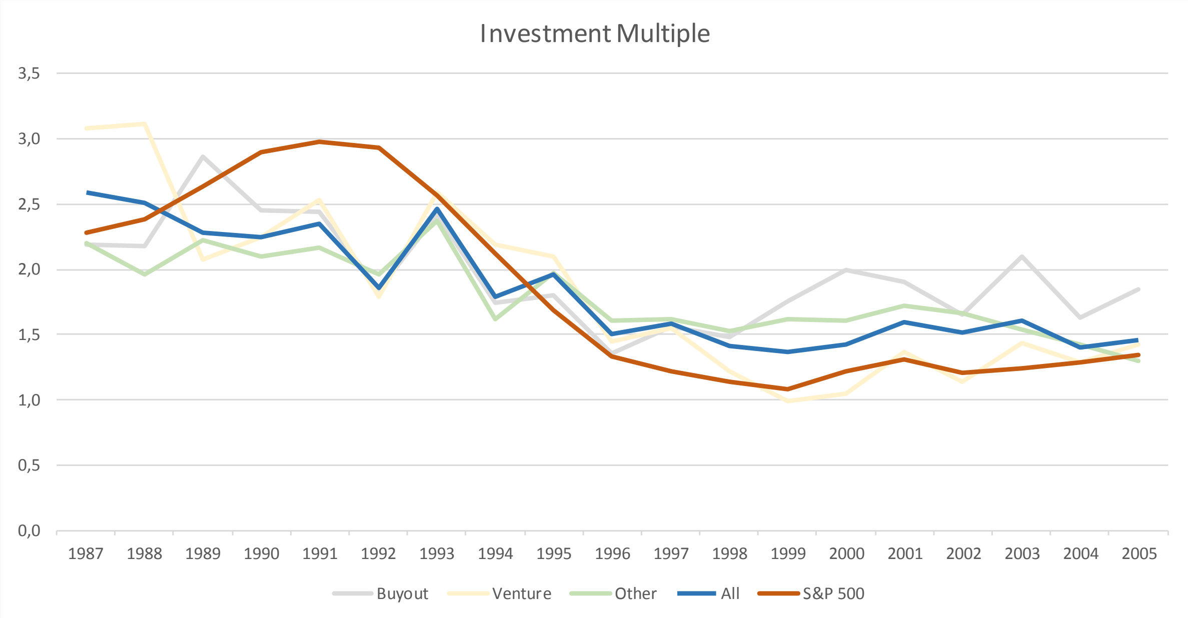
The biggest and most obvious drawback of using multiples, is that they do not take into consideration the timing of capital calls and distributions, nor does it take into consideration time value of money. Even though these metrics are relatively easy to understand, without the time dimension, one could get the same results by putting money in the bank and waiting. Therefore, time dimension is a critical factor when comparing actual fund performance.

If multiples are to be used, they should be accompanied by cash flow data as well as forecasts on when capital calls and distributions will occur. These forecasts would be difficult to produce, and even more so for young funds, and funds run by inexperienced GPs. Such a forecast would thus be unreliable.

Based on these drawbacks, multiple comparisons need to rely on fairly strong and general assumptions regarding calls and distributions. These assumptions are too stringent to be used in comparisons of funds on a fund-by-fund basis. However, they could be used in comparing different investment strategies, like industry focus, region focus or type of fund. In this setting, it is possible to make more general assumptions regarding PE cash flow cycles, since they on average follow similar pay-in and distribution cycles. From this, it is possible to construct an equivalent public market investment vehicle and compare strategies to this portfolio.

4.1.2.1 Peer Group Comparison

Example



Fund Cycle	1st year	2nd year	3rd year	4th year	5th year	6th year	7th year	8th year	9th year	10th year
Weights	10 %	40 %	70 %	90 %	90 %	90 %	90 %	60 %	40 %	20 %

Figure 1 - Comparing PE Multiples and an S&P500 Investment Vehicle

In the above graph, we have compared average TVPIs to the return LPs would have gotten if they used a similar S&P500 investment vehicle. These weights emulates the percentage of committed capital normally tied up in PE investments.

$$S\&P\ vehicle\ multiple = \prod_{t=1}^{T} 1 + (Anual\ S\&P\ return_t \times Weight_t) \tag{4}$$

From the graph, we see that PE in general outperforms public equity, except in the period from 1989 to 1993. This is a broad statement, and is sensitive to the chosen weights of the S&P500

investment vehicle. Therefore the amount of under or over performance can be altered drastically. However, the trend is less affected by changes in weights and we can see that there is some correlation between PE returns and public equity returns. This correlation is arguably not causal. Both returns probably rely on some other unobserved factor that influences the return of both PE investments and public stock market investments.

Since the norm in evaluating PE funds is peer group comparisons, these peer groups can take on any form an LP chooses. They will in general include vintage, fund type and area of focus (either geographic, industry or both). PE data providers will often let investors create custom benchmarks in order to compare performance of funds that are in line with their own investment strategy more accurately.

Because peer group comparisons are the norm, a comparison of buyout, venture or other types of PE funds, should be compared to a public index that more closely resembles the types of companies a fund is likely to invest in. It would therefore be better to use an index like Nasdaq Small-Cap or Russell 2000® to compare venture returns to public market returns. A comparison with the Dow Jones Large-Cap index or the MSCI USA Large-Cap index might be a better basis for comparing buyout returns to that of public equity.

4.1.3 Internal Rate of Return

IRR is the most widely used PE performance metric. It is also used in the evaluation of other forms of corporate investments. LPs and other corporate investors are familiar with this performance metric, and this might be a part of the reason for its widespread success. The IRR also facilitates easy comparison between investing in PE and investing in other corporate projects, however, it is not easily comparable to the returns gained from public equity investing (Kaplan & Schoar, 2005).

Another important reason for the success of IRR is that it, in contrast to multiples, takes into consideration the timing of cash flows. In its theoretical form, IRR is the discount rate ensuring that the net present value of cash flows is zero. The GIPS (CFA Institute, 2010) propose this calculation of interim return measurement:

$$0 = \sum_{i=0}^{n} CF_i \left(1 + \frac{r}{c}\right)^{-(ic)} \tag{5}$$

Where CF is the cash flow for period i, n is the total number of cash flows, i is the cashflow period, c is the number of annual cash flow sub periods, and r is the sub period IRR.

The IRR favours early cash flows and thus hinders GPs accumulating capital at the beginning of a fund's life. It also incentivises GPs to distribute proceeds quickly after they have been realised. There are numerous pitfalls when comparing fund performance based on IRR, some of which are also evident in evaluating corporate projects. LPs need to be aware of these before an investment decision is made. Because of these pitfalls IRR has been criticised by a number of papers (Higson & Stucke, 2012; Phalippou, 2008), and the main pitfalls are outlined below.

4.1.3.1 Aggregation issues

A problem with using IRR is that the average is different from the aggregated cash flows. This can potentially be a big problem in comparing PE returns since a fund's IRR is negatively related to duration, meaning the average performance is usually upwardly biased. Difficulties may arise when comparing funds based on an industry average, or by other characteristics like fund type or size. Because of the duration issues, funds with longer duration will usually underperform based on an average IRR comparison.

There might be underlying factors that results in some fund types having consistently lower duration than others. This could lead us to wrongly conclude that they outperform other types of funds. Phalippou and Gottschalg (2008) suggest that a weighting based on duration might be a step in the right direction. However, this requires cash flow data. A duration weighting seems like an intuitive correction and means that funds with different timing of cash flows will be treated differently.

If cash flow data is not available, we need another way of detecting differences in timing of cash flows. A comparison of TVPI and IRR is therefore used in our thesis. We would expect TVPI and IRR to behave similarly if the cash flows on average have similar durations and timing of calls and distributions.

4.1.3.2 Endogenous Cash Flows

The problem with endogenous cash flows is that it provides GPs with incentives to strategically time calls and distributions. By waiting to draw down capital from LPs, as opposed to requiring payment upfront, GPs are able to minimise the time element, and therefore allows them to maximise IRR. Thus, GPs have the ability to game their cash flows.

Since IRR is biased, and favours early cash flows, this incentivises GPs to get out of good investments early, and hold on to bad investments longer.

Buyout funds have been criticised for buying a company, borrow large amounts of capital with the company's assets as collateral, and using the borrowed money to pay out large dividends. Another criticised practice that is quite common is to take a company public and distribute shares directly to LPs. Both of these practises are in line whit the attempt to maximize IRRs (Hall, 2006).

Although there is now proof that inflating IRR is the reason behind these practices, buyout firms have been called "evil empires". In the 1980s, managers like T. Boone Pickens and Carl Icahn became infamous for buying companies and streamlining production by selling of large amounts of assets to increase exit multiples, and hence a company's valuation (Cendrowski, Petro, Martin, & Wadecki, 2012, p. 165).

4.1.3.3 Reinvestment Assumption

The IRR equals the effective rate of return only if intermediate cash flows distributed by the PE fund can be reinvested in other opportunities at the same rate. If the IRR is high, the spread between IRR and effective rate of return is positive and large. If the IRR is low, the spread is negative and large. Concequently, funds with a high IRR have an IRR greater than the effective rate of return, with the opposite being true for funds with a low IRR.

Based on this, and the volatility of intermediate cash flows, IRR can be misleading. Results therefore show a more dispersed performance figure than what might actually the case.

4.1.3.4 Valuation risk

During a fund's life, IRRs are calculated by taking into account the unrealised value of investments. This is in line with the calculations of RVPI in the PE multiples case, which means that the interim estimates of IRR must be based on expected future cash flows. The GPs consequently have a potential to manipulate results and overstate expected returns on exits. This problem is reduced as a fund matures, partly because GPs have more information regarding their own portfolio companies, and partly because the closer a fund gets to liquidation, the less impact cash flows have.

To take the unrealised value of investments into consideration, NAV is used. The NAV is mostly used for public companies and is the assets less liabilities, divided by outstanding shares. In the case of PE, the expected present discounted sum of future cash flows is one way NAV is calculated. The British Venture Capital association, one of the associations that helped form the GIPS guidelines, presents this method of calculating NAV (GIPS, 2006; Steer & Ellis, 2011).

$$NAV_i = E_i \sum_{i=j}^{N} \beta^i \, CF_i \tag{6}$$

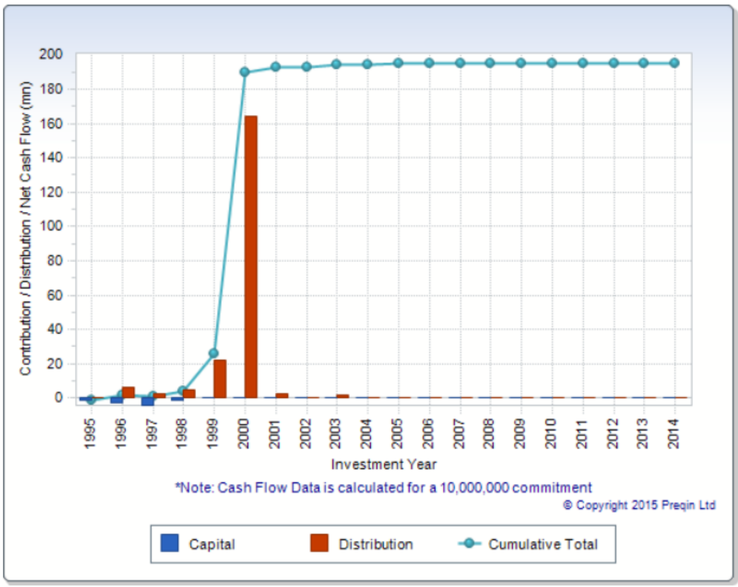
Where β is the discount rate, i is the period and N is the number of periods.

Because there is some subjectivity present in estimating both the discount factor and the future cash flows, uncertainty around interim IRRs for firms that are not yet liquidated arises. This also makes it possible to, either deliberately, or by accident, over or understate NAV.

4.1.4 Real World Cash Flow, IRR and TVPI Example

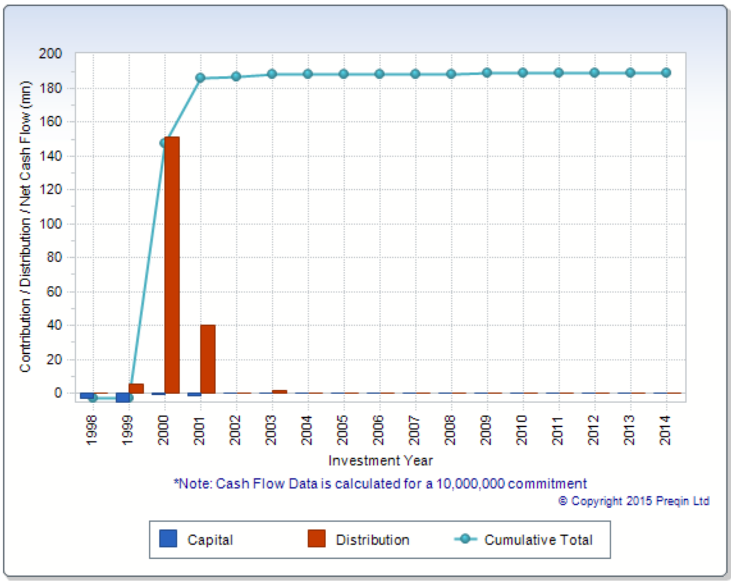
An example of the potential weakness of IRR is the case of Example Partners and their fund I and II[2]. The funds show IRRs of 218.3% and 514.3%, respectively. A closer look at the funds' cash flow data, reveal huge distributions in the funds' early years, with marginal distributions later in time. This gives the funds an effective lifespan of 11 years, however, most distributions happen in the early years.

[2] The funds in this example are real, and are collected from Preqin's database, but due to confidentiality, the names have been changed.



Example Partners I - Cash Flow Data			
Vintage	1995 Size		125
IRR			218,3
TVPI			20,51
Date	Called %	DPI %	IRR
1995	15,0	0	n/m
1996	45,0	145,16	n/m
1997	84,7	108,65	n/a
1998	100,0	137,33	n/a
1999	100,0	358,01	214,6
2000	100,0	1997,17	218,4
2001	100,0	2024,18	218,3
2002	100,0	2024,19	218,3
2003	100,0	2045,13	218,3
2004	100,0	2015,13	218,3
2005	100,0	2045,13	218,3
2006	100,0	2051,10	218,3

Figure 2 – Example Partners I – Cash Flow Illustration



Example Partners II - Cash Flow Data			
Vintage	1998 Size		200
IRR			514,3
TVPI			19,86
Date	Called %	DPI %	IRR
1998	30,1	0	n/m
1999	80,0	65,39	n/m
2000	87,5	1786,50	n/a
2001	100,0	1961,68	515,6
2002	100,0	1964,20	515,0
2003	100,0	1964,20	515,0
2004	100,0	1982,45	515,0
2005	100,0	1982,45	515,0
2006	100,0	1982,45	515,0
2007	100,0	1982,45	515,0
2008	100,0	1982,45	515,0
2009	100,0	1986,22	514,3

Figure 3 – Example Partners II – Cash Flow Illustration

There is no denying that Example Partners has been extremely successful, returning around 20 times the initial investments of LPs in both funds. What is evident from the cash flows is that both funds return roughly the same multiple, but because of the timing of distributions, fund I displays an IRR that is less than half of fund II. Both funds have similar distributions from 1998 and onwards with the bulk being distributed before the collapse of the dot-com bubble. Because fund I started calling on capital three years earlier than fund II, the IRR is substantially smaller, and clearly shows that IRR is negatively correlated with a funds duration.

In the case of aggregation issues, both funds have a much larger IRR and TVPI than the other funds with the same vintage. These funds will skew the average IRR severely if such a measure is used. If cash flow data is available, a duration weighted benchmark, as proposed by Phalippou and Gottschalg (2008) would be a better measure.

If cash flow data is not available, the use of median IRRs or an average, which excludes extreme values, will also mitigate the outlier problem. In the statistical analyses, we have used one dataset based on median IRRs, and another dataset based on mean IRR excluding extreme values.

5 Data

This thesis is based on a dataset from Preqin. Preqin gets their data in several different ways, the main sources being GPs, LPs and Freedom of Information Act (FOIA) requests. Also, regulatory filings and monitoring of media outlets provide useful data. Direct correspondence with Preqin tells that GPs provide 60% of the performance data (Harris et al., 2010). FOIA is mostly used if fiduciary responsibilities do not allow for disclosures.

Preqin has been researching the PE industry for over a decade. According to our Preqin contact, the data is trusted by the most respected alternative asset media outlets like Bloomberg, Financial Times and Wall Street Journal amongst others (J. Kimble, personal communication, May 13, 2015). These media outlets are known to have the highest quality data on the market and are heavily relied on by the largest global banks, fund managers, investors and law firms. As of the 1st of May 2015, Preqin covers 20 448 PE firms, 43 073 funds, 19 995 funds with performance data, 6 004 funds with IRR data and claims to have the best market coverage (Preqin, 2015).

Harris, Jenkinson, & Stucke (2010) point out that GPs may not be incentivised to provide IRR, but Preqin themselves claims to have the best net to LP performance data (Preqin, 2008). However, when research relies on voluntary submission of data, there could be a problem with survivorship and backfill bias. Survivorship bias occurs when poor performing funds stop reporting results and falls out of calculations. Backfill bias occur when funds stall their performance reporting only to backfill them when better results have been achieved. According to Russel (Gupta, 2012) and Preqin, there seem to be no survivorship bias in Preqin’s dataset, but Harris, Jenkinson, & Stucke state that it could suffer from backfill bias.

The reliability of IRR is often questioned when analysing performance data. We will discuss the benefits and disadvantages of IRR later, but we would like to quote what Preqin had to say about their own IRR calculations:

> *The IRR is extremely reliable whether the fund is liquidated or not. When we calculate it ourselves we use the cash flow data to get an accurate calculation. For the firms that just report IRR, we not only benchmark them against their industry to ensure performance is in line but we also contact investors to make sure we are getting accurate information (J. Kimble, personal communication, May 13, 2015).*

5.1 Data Processing

The original dataset contained records of 22 048 PE funds. Many of these funds had missing data and editing needed to be done.

First, the sizes of the funds were in nominal terms, so we adjusted them for inflation (Bureau of Labor statistics, 2015).All fund sizes are now presented in 2008 dollars.

Second, we limited our data to funds with vintage between 1990 and 2008. There are few observations in the dataset before the 90s, so in order to get proper measures for decades, we chose to start at 1990. The reasoning behind the 2008 cut-off relies on a few factors. Possibly, the best data would come from already liquidated funds. Using liquidated funds may be more reliable as the numbers going forward are actually realised. However, by only accepting liquidated funds into the dataset, we would have reduced the dataset by approximately ⅔, leaving us with too few observations to make any meaningful inferences. Ljungqvist & Richardson (2003) and Steer & Ellis (2011) find that there are no systematic bias when a fund is sufficiently mature, hence we include observations up until 2008.

Third, we only kept funds with both focus and GP location in the US. By doing this, we avoid possible problems like difference in legislation or other governing factors between countries and regions.

Fourth, we dropped all the funds that either had missing data for IRR, size or TVPI. We assume that the data missing is not due to some underlying characteristics and therefore dropping them will not create any bias.

Fifth, only buyout and venture funds were kept. There were many different fund types in the original sample, but due too few observations for all but buyout, venture and real estate, these were dropped. Most PE research have focused on either buyout, venture or both of them, making this study more comparable to previous work. We have also found that there is no significant differences in distributions between the performance of real estate and the performance of buyout and venture combined. Hence, real estate was dropped, too.

After having dropped the necessary data, a sample of 786 PE funds were left. In the analysis later on, we are going to run two different tests depending on different sample characteristics. The student’s t-test for two independent samples (t-test) is used to compares means, while the

Mann-Whitney Wilcoxon test (MWW-test) is used to compares medians. When comparing medians, the sample containing 786 funds is sufficient, and we call this sample the *untrimmed dataset*. However, when comparing means, the results tend to be upwardly biased because of aggregation issues. We reduce this problem by cutting the top and bottom 2.5% for IRR, size and TVPI. By doing this we drop 118 observations, ending up with a sample of 668 PE funds. We call this sample the *trimmed dataset*. As we will use the t-test in most of the analysis, the descriptive statistics will rely on the trimmed dataset. For descriptive statistics on the untrimmed dataset, see Table 6, Table 7 and Table 8 in Appendix B – Descriptive Statistics for Untrimmed Dataset.

5.2 Descriptive Statistics

The mean[3] IRR (size)[4] [TVPI][5] of the sample is 9.73% ($403M) [1.56]. Table 3,Table 4 andTable 5 in Appendix A show the whole descriptive statistics for the trimmed dataset for IRR, size and TVPI, respectively. As seen in these tables, we control for some specific factors, including fund types, cycles, decades, sizes and sequence numbers.

5.2.1 Fund Types

The first factor we control for is fund types. As mentioned, only buyout and venture funds are present in the sample. There are 271 buyout funds and 397 venture funds, making the sample fairly well distributed. Buyout (venture) funds have an IRR of 13.32% (7.27%), a size of $545M ($306M) and a TVPI of 1.77 (1.42).

5.2.2 Cycles

Secondly, we control for business cycles. We have divided the business cycle into boom and bust periods. Some papers control for these periods, but few, if any, mention the specific time periods of these cycles. Kaplan & Strömberg (2008) are speaking of buyout booms in the late 80s, early 90s and between 2005 and mid-2007. They also state that a boom can only happen when earnings yield (S&P 500 companies in this case) exceeds interest rates on high-yield

---

[3] Unless specified otherwise, future mentions of numbers relating to IRR, size or TVPI will always be in mean.
[4] Unless specified otherwise, all numbers mentioned in parentheses during the rest of this section will be size numbers
[5] Unless specified otherwise, all numbers mentioned in brackets during the rest of section 4 will be TVPI numbers

bonds. This is not sufficient though, and other condition also need to be met in order to experience a boom. Acharya, Franks, & Servaes (2007) speak about boom and bust periods, too, but they only look at buyout as well. According to them, the buyout boom in the 00s lasted from 2001-2006. However, none of them explain the criteria for defining these periods.

We classify each individual year in the sample as either boom or bust, where bust is a year containing at least six months of a recession. Looking at the recessions between 1990-2008 (the National Bureau of Economic Research, 2015), 1990, 2001 and 2008 are considered bust periods in our sample. The rest is classified as boom. Since the records only keep track of funds' vintages and not the specific dates they are raised, this is about as precise as the classification can get. A fund started in January 1991 is in principal started during a recession, but as the recession ended in March 1991, this year has been classified as boom, and thus the fund is classified as boom, too. Some funds will therefore have similar characteristics, but will be classified differently. Counting observations, boom and bust have 567 and 101, respectively. Hence, problems will arise later on, concerning too few observations in bust periods. When controlling for sequence numbers and bust simultaneously, the possibility of getting insignificant results improves. The IRR of boom funds is 9.18% ($406M) [1.53], while the IRR for bust funds is 12.76% ($386M) [1.71].

5.2.3 Decades

Thirdly, we control for decades. As the sample stretches from 1990-2008, we classify them as either 90s (1990-1999) or 00s (2000-2008). We have an overweight of funds in the 00s, counting 376 observations in this decade compared to 292 in the 90s. The 90s have an IRR of 12.11% ($390M) [1.65] while the 00s have an IRR of 7.87% ($413M) [1.49].

5.2.4 Size

Fourthly, we control for size. To do this, we classify all the funds smaller than $100M as small. Funds equal to or larger than $100M, but smaller than $500M, are classified as medium, while funds with a size of $500M and above, are classified as large. By doing this, small funds will be dominated by venture while buyout will dominate large. This is much due to the nature of these fund types, as discussed earlier. Looking at observations, small, medium and large count 127, 360 and 181, respectively. The IRR of small funds is 10.68% ($58M) [1.59] while the equivalent measure for medium and large is 9.85% ($259M) [1.59] and 8.82% ($931M) [1.47], respectively.

5.2.5 Sequence Numbers

Last, we control for sequence numbers. We have divided the funds into classifications as shown in Table 1.

Sequence Number (SN) Classifications	
SN	Explanation
0&1	A firm's first fund
0	A firm's first fund, but no follow-on fund has been created
1	A firm's first fund, and at least one follow-on fund has been created
2	A firm's second fund
3	A firm's third fund
4	A firm's fourth fund or more

Table 1 - Sequence Number Classifications

The classification of sequence numbers was done before any data was dropped. To show why we did this and how the classification works, we will use an example containing the imaginary funds in Table 2.

Example

As we can see, Imag PE Partners started their PE business in 1986. In 1991 they created their second fund (Imag Buyout I), which meant that the first fund had a follow-on fund. Hence, Imag Venture I was classified as 1 and not 0. Also, two more funds were created in 1991, Imag Venture II and Imag Buyout II. These are also classified as sequence number 2. In our opinion, there is one upside and one downside to this. The downside is that Imag Buyout I and II will both have the same sequence number. Although we do not have the exact dates these funds were raised, it seems obvious that Imag Buyout I was created before Imag Buyout II. The upside is that we may capture more of the sequence number characteristics this way. GPs tend, on average, to create a fund every three years. If a GP's sequence number 2, 3 and 4 were created in the same year, it would be difficult to capture size effects. LPs who invest in Imag Buyout II will not be aware of the performance of Imag Buyout I. We look at this upside as bigger than the downside, and hence classify all funds created in the same year with the same sequence number. Finally, the two last funds of Imag PE Partners are both sequence number 4, as sequence number 4 contains a GP's sequence number 4 or above. This is why ⅓ of the observations belong to sequence number 4. For descriptive statistics on sequence numbers, see Appendix A.

Imag PE Partners Fund History				
Firm Name	Fund Name	Vintage	Sequence Number	In sample
Imag PE Partners	Imag Venture I	1986	1	No
Imag PE Partners	Imag Buyout I	1991	2	Yes
Imag PE Partners	Imag Venture II	1991	2	Yes
Imag PE Partners	Imag Buyout II	1991	2	Yes
Imag PE Partners	Imag Venture III	1995	3	Yes
Imag PE Partners	Imag Venture IV	2003	4	Yes
Imag PE Partners	Imag Venture V	2009	4	No

Table 2 - Classification of Sequence Number

5.3 Quartile Data

Past GP performance is widely used by LPs when picking funds to invest in. A fund's performance is therefore often accompanied by its quartile rank. This rank is established by comparing the fund's IRR with the IRR of similar funds. Preqin's default metric for computing the benchmark IRR, is a median of funds from the same vintage, same fund type and funds focusing on the same location or region. These characteristics cannot be upheld in all cases. Depending on the information available, the number of funds in a peer group or an investor's preference, these can be changed to better reflect performance, and to make portfolio comparisons possible.

In our dataset, we have included Preqin's default benchmarks. We then looked at GPs which had funds in a previous vintage, and linked the performance data of the previous fund to the next fund they raised.

In 2014, Preqin published a press release announcing the most consistent performing GPs (Preqin, 2014). They looked at the last three funds a GP had, which had a similar investment strategy. Preqin used their own quartile ranks based on both TVPI and IRR. This should make gaming of quartile rank more difficult and therefore make the rankings more robust than those of for example VE.

6 Hypotheses

In the analysis, we will compare fund characteristics against each other to check if some of them are significantly different. The hypotheses are based on previous research, and our own subjective opinions.

6.1 IRR and TVPI

Both IRR and TVPI are performance metrics and should not show very different results, unless there are some other underlying factors that need special attention. Hence, our hypotheses for these measures are equal.

6.1.1 Sequence Numbers

Sequence number 0 only contains funds with no follow-on funds. Hence, we expect them to perform worse than funds with higher sequence numbers. This coincides with the finding of Chung et al. (2012), that high performing funds are more likely to have a follow-on fund. Harris, Jenkinson & Stucke (2014) also find that well-performing funds, in most cases, seem to raise new, well-performing funds. There may of course be several reasons, but we do believe that performance is an important decision factor when considering raising a follow-on fund.

The same arguments holds for sequence number 1. To create a follow-on fund, the first fund usually performs well. Hence, we believe that sequence number 1 outperforms all other sequence numbers.

Looking at sequence number 0&1, our view depends on the number of observations in sequence number 0 and 1. An overweight of observations in sequence number 0 indicates that few follow-on funds are raised. Hence, we believe performance among first-time funds are poor. However, should there be an overweight of sequence number 1, we believe the opposite will happen. Comparing 0&1 against sequence number 2, 3 and 4, Kaplan & Schoar (2005) find that first-time funds perform worse than funds with higher sequence numbers. A factor pulling in the other direction is the experience and skills of those GPs that has managed to raise follow-on funds. Based on theory and our own opinion, we expect that the lack of experience will make first-time funds underperform.

Comparing sequence number 2, 3 and 4, we expect there to be an upwardly trend due to increased experience, and the fact that poor-performing GPs will not be able to raise follow-on funds. This will in turn weed out poor performers, and we should be left with a higher share of skilled GPs managing funds with higher sequence numbers.

6.1.2 Fund Characteristics

6.1.2.1 Fund types

To compare buyout and venture funds, we need to take a closer look at their investments. The most notable difference between them is the characteristics of the companies invested in. In the buyout industry, portfolio companies are often well established, while the venture industry is packed with young companies and entrepreneurs looking to enter the markets. We believe there are greater risk involved in the venture industry, as far from all venture-backed companies succeed in their pursuit of success. Thus, venture funds might experience more cyclical returns, and in our opinion underperform compared to buyout funds.

6.1.2.2 Cycles

Looking at cycles, boom periods are much longer than bust periods. Returns in general are higher during boom periods, enabling funds that are active in more years of high economic growth, to gain higher returns. In addition, during bust periods, prices tend to fall. Hence, funds raised in bust periods make their investments at lower prices, increasing chances of greater returns. Obviously, funds raised at the start of a boom period will experience much of the same effects and have many of the same characteristics. This could possibly reduce the differences between periods. However, we still believe that funds raised in bust periods will perform better than funds raised in boom periods.

6.1.2.3 Decades

Comparing decades, we look at key events during the 90s and the 00s. As mentioned earlier, reporting of PE performance have improved over the last decades. If there is systematic overstatement of returns in the 90s, we believe there is an upward bias in the 90s compared to funds raised in the 00s. Due to cheaper financing and a general growth in PE during the 00s, the industry has been more accessible in recent years. An increase in demand from LPs could potentially lead to more funds being raised just to fill this demand, favouring quantity over quality. Hence, we expect returns to be higher in the 90s.

The recessions from the 1990 until 2008, play a big part in explaining the differences between fund types during different decades. Leading up to the burst of the dot-com bubble in 2001, most venture-backed companies experienced enormous growth. We would at least expect venture funds raised in the early 90s to perform well, and therefore outperform buyout funds in this decade. While the venture industry took some time to recover after the recession, the buyout industry benefited from a long buyout boom in the 00s. We thus believe that buyout outperform venture in the 00s. Both the long boom period during the 90s and the buyout boom in the 00s, have contributed in generating good returns for buyout funds. Hence, we find it hard to expect differences in performance across decades. Due to the dot-com bubble in the late 90s and early 00s, we expect venture funds raised in the 90s to outperform venture funds raised in the 00s.

6.1.2.4 Size

Gompers & Lerner (2000) find that larger firms seem to pay a higher price for their investments. By being large, it may be easy to grasp over too much, being less concerned about the price of an investment. When comparing fund sizes, this implies that smaller firms have better performance. Also, Kaplan & Schoar (2005) find that successful GPs chose not to grow as much as less successful. However, larger funds have a greater possibility to diversify their investments, reducing the amount of unsystematic risk. Although we do not think differences in size influence performance too much, we favour smaller funds over larger.

Metrick & Yasuda (2010) find that buyout funds are more scalable than venture funds. This is mainly due to advantages concerning debt financing, but we also believe there is another reason. Controlling venture-backed companies demand huge resources, mainly human skills. By investing in too many companies, GPs would not be able to use the necessary amount of time and dedication to fulfil each investment's potential. Hence, we believe most venture funds perform better when smaller. Looking at buyout funds, Metrick & Yasuda find that some buyout funds expand rapidly, favouring short-term income from fees over quality. This points towards better performing, smaller buyout funds. Comparing fund types on sizes, we would expect insignificant differences among small funds, but expect the differences to increase with larger funds.

6.2 Size

6.2.1 Sequence Numbers

Looking at sequence numbers, we expect performance and time to be the main factors affecting size. First, well-performing GPs tend to attract more investors, increasing fund sizes as they raise follow-on funds. This implies that fund size increases with sequence number. However, if a follow-on fund is raised only a year or two after the first, investors may not be able to see how the first is performing. Hendershott (2008) also points out that a fund needs at least four years to be able to predict, with 50% certainty, that a fund with interim top quartile performance will finish in the top quartile. Hence, a follow-on fund may not be larger than its predecessor.

Second, we have seen a general growth in the economy, and expect the PE industry to follow the same path. Given that the industry grows faster than the inflation, we expect fund size to be positively correlated with sequence number. Thus, our hypothesis is that size increases by each sequence number. Sequence number 0 and 1 are exceptions here, as they both are a GP's first fund and we expect them to be the same size.

6.2.2 Fund Characteristics

6.2.2.1 Fund Types

We mentioned that buyout funds are more scalable than venture funds, implying they might be larger. Since we also expect venture funds to benefit from being smaller, our prediction is that buyout funds in general are larger than venture funds.

6.2.2.2 Cycles

The size of funds depend heavily on when they are raised. LPs may be less willing to invest money in bust periods as investors usually become more cautious during recessions. Thus, we believe that funds raised in boom periods are larger.

6.2.2.3 Decades

The PE industry has evolved over the last two decades. With a bigger interest in the asset class now than earlier, more capital is being invested, possibly leading to larger funds in recent years. However, the 00s have experienced two bust periods, implying funds in the 00s may not be that large. We believe the former argument is stronger, though, thus expecting funds raised in the 00s to be larger than funds raised in the 90s.

Looking at fund types, the venture boom during the 90s would imply better venture performance. As performance increases, demand increases, possibly increasing fund sizes. However, we do believe that buyout funds are large due to the scalability of the fund type. Hence, we expect small differences in sizes during the 90s.

Although most of the 90s was a period of high economic growth, there was a big buyout boom in the 00s. Combined with the growth in the PE industry over time, we believe buyout funds will be larger in the 00s compared to the 90s. The strong venture performance in the 90s makes us believe that venture funds are larger in this decade.

6.3 Quartile and Past Performance Persistence

There is a possibility that a GP in the top quartile in one period got there because of luck. Also, a proportion of skilled GPs will have their funds outside of the top quartile because of bad luck. We would therefore expect, looking one period back at a GPs' last fund quartile performance, it is not a very significant indicator of the current fund's performance. However, we expect that this becomes more significant if we look back several periods.

Robert Hendershott (2008) has suggested that GPs need three or four previous funds in the top quartile to be able to predict top quartile performance for their next fund. Hendershot used VE as a data source, and although this dataset has been found to exhibit weaknesses, we expect the results from Preqin's database to give similar results. Our expectations are therefore that a previous period's quartile rank is not a significant indicator of next fund's performance.

7 Methodology

7.1 Statistical tests

To be able to perform the analyses done in this thesis, we have run a series of tests using both Microsoft Excel and Stata. The main test used is the t-test, but we also make use of the MWW-test.

7.1.1 Student's t-test for two independent samples

The t-test is used to determine if there is a statistically significant difference between the means of two independent groups on a continuous dependent variable. The dependent variable will in this thesis be either IRR, size or TVPI. All these variable are measured at a continuous level, even though one could argue that the upper range of these variables is in fact infinitely.

There are two possible formulas that can be used when computing the t-test. One is run if we assume equal sample variances (Formula XXX), and another is run if the variances are unequal (Formula XXX). The t-test assuming equal variances is given by

$$t = \frac{\bar{x} - \bar{y}}{\left(\frac{(n_x - 1) \times s_x^2 + \left(n_y - 1\right) \times s_y^2}{n_x + n_y - 2}\right)^{1/2} \left(\frac{1}{n_x} + \frac{1}{n_y}\right)^{1/2}} \tag{7}$$

and the t-test assuming unequal variances is given by

$$t = \frac{\bar{x} - \bar{y}}{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}\right)^{1/2}} \tag{8}$$

where t is the test score, $\bar{x}$ and $\bar{y}$ are the averages of the first and second sample, respectively, n_x and n_y are the number of observations for the first and second sample, and s_x^2 and s_y^2 are the variances for the first and second sample.

If the test is significant, we reject the null hypothesis of equal population means in favour of the alternative hypothesis of difference in population means. In order to perform the t-test, five different assumptions need to be met. There needs to be:

1. One independent variable consisting of two categorical, independent groups.
2. Independence of observations.
3. No significant outliers in the two groups in terms of the dependent variable.
4. Approximate normal distribution for the dependent variable for each of the two groups.
5. Homogeneity of the variances for the two groups.

If the third assumption is violated, it could have a large influence on the mean and standard deviation for the group, thus affecting the test results. The outlier problem is more severe if a sample size is small. To mitigate this problem we have trimmed the dataset to remove extreme values for IRR, size and TVPI. Whenever the t-test is applied to our analysis, we only used the trimmed dataset.

Even though the fourth assumption should be upheld in order to get proper test results, the t-test is somewhat robust to violations. Therefore, the data only needs to be approximately normally distributed, and because of the central limit theorem, the validity of the results increases as the sample size increases. If a sample size is small, the t-distribution is a poor approximation to the t-statistic if we are dealing with non-normality. However, as sample size increases, the estimator will satisfy asymptotic normality. There are no general consensus as to how big a sample size must be before the approximation is good enough, but a general rule of thumb is 30 observations (Wooldridge, 2014).

A potentially bigger problem than a non-normal distribution, is non-homogeneous skewness in the distributions of the two samples. A violation of the fifth assumption becomes more severe the bigger the difference is between the sample sizes in each group. If the sample sizes are similar, a violation is often not that serious.

7.1.2 Mann-Whitney Wilcoxon

The MWW-test is a non-parametric alternative to the t-test and is more efficient if the data is non-normal. If we find, after evaluating the data, that neither performance metric is very normal nor particularly symmetric, the MWW-test will be a more appropriate way of comparing different characteristics in our data. The first two assumptions for the t-test also applies to the MWW-test. The test is used to check whether there are differences in the distributions of two groups, or differences in the medians of two groups, and is given by

$$z = \frac{\overline{U} - min(U_x, U_y)}{\sqrt{Var(\overline{U})}} \tag{9}$$

where z is the test score, $\overline{U}$is the mean U-score for the two samples and $\min(U)$ is the lowest of the two U-scores for the two samples. U is given by

$$U_{x,y} = n_{x,y} \times n_{y,x} + \frac{n_{x,y} \times (n_{x,y} + 1)}{2} - R_{x,y} \tag{10}$$

where U is the U-score, n is the number of observations in a sample and R is the rank sum for a sample. R is the sum of all the ranks for a given sample.

The MWW-test works by ranking each score of the dependent variable according to size, and without consideration to which group it is in. The ranks obtained for the two samples are then averaged and tested for differences. A numerical example is given in Appendix I.

As mentioned, the MWW-test interpret whether there is a difference in the distributions of two groups or if there is a difference in the medians of two groups. Which test is carried out is dependent on the distribution of scores for both groups of the independent variable. This in turn leads to two possible alternative hypotheses:

HA1: The medians of the two groups are unequal

HA2: The mean ranks of the two groups are unequal

If we consider the first alternative hypothesis, we are testing if the medians are different for the two groups. This is more in line with the t-test, which does the same for sample means, and assumes that the shape and dispersion of the distributions are similar. In presented tables, this type of analysis will be indicated by subscript 1 (MWW_1).

If we consider the second alternative hypothesis, we are testing for differences in distributions. We would here be interested in whether the performance of sample 1 and sample 2 are similar, or if one sample is significantly different from the other. The MWW-test does this by comparing the mean ranks of each distribution. In presented tables, this type of analysis will be indicated by subscript 2 (MWW_2).

7.2 Test Selection Process

Below is a flowchart that describes the selection process used for deciding between the parametric t-test and the non-parametric MWW-test. We start by using the trimmed dataset, as this is the basis for the t-test. The first thing we need to check is whether the samples are normally distributed or not. To do this, we use the skewness and kurtosis test for normality (SK-test). The SK-test is given by

$$K^2 = Z_1^2 + Z_2^2 \tag{11}$$

Where K^2 is the test score and Z_1^2 and Z_2^2 is the distribution of the test statistic for skewness and kurtosis, respectively. K^2 has an approximately x^2 distribution with two degrees of freedom. If K^2 is significant, the null hypothesis of normal distribution is rejected in favour of the alternative hypothesis of a non-normal distribution. In the case where both samples are normally distributed, the t-test will be used. However, before the t-test is run, we need to check whether the variances of the two samples are equal or not. This is done using the F-test for two samples variances (sdtest), which is given by

$$F = \frac{s_x^2}{s_y^2} \tag{12}$$

where F is the test score and s_x^2 and s_y^2 are the variances of the first and second sample, respectively. If F is significant, the null hypothesis of equal variances is rejected in favour of the alternative hypothesis of unequal variances.

Should the samples show non-normal distributions, we need to check the number of observations in each sample. As mentioned earlier, the t-distribution is a poor approximation to the t-statistic if we are dealing with non-normality and a small sample size. We choose to use 50 observations as a lower limit, and all samples with less observations are generally ruled out for the t-test. However, we do study all samples carefully. Should some of those samples with less than 50 observations show very clear signs of being eligible for the t-test through our skewness tests, the t-test will be used.

We use three different ways of looking at skewness for samples with more than 50 observations, determining whether they are fit for the t-test or not, even though the samples are not normally distributed. Firstly, we see if the distributions are similarly skewed, simply by looking at the skewness coefficient. Secondly, we use the SK-test to determine if both distributions are significantly skewed. Thirdly, we take a closer look at the distributions (graphically) to double check that they show the same as the two former tests. When deciding if the samples are fit to be used in the t-test, some subjective choices are made. We try as best we can to make the right calls, but there is always the possibility that two samples should have been compared using the MWW-test instead. If the data fails to meet the criteria of the t-test, we will use the MWW-test and switch to the untrimmed dataset.

The final step of the process is to see if the samples have a similar shape. If they are similar, we can make an inference based on the medians of the two samples, whereas if they are dissimilar, we are limited to make an inference based on the distributions of the two samples.

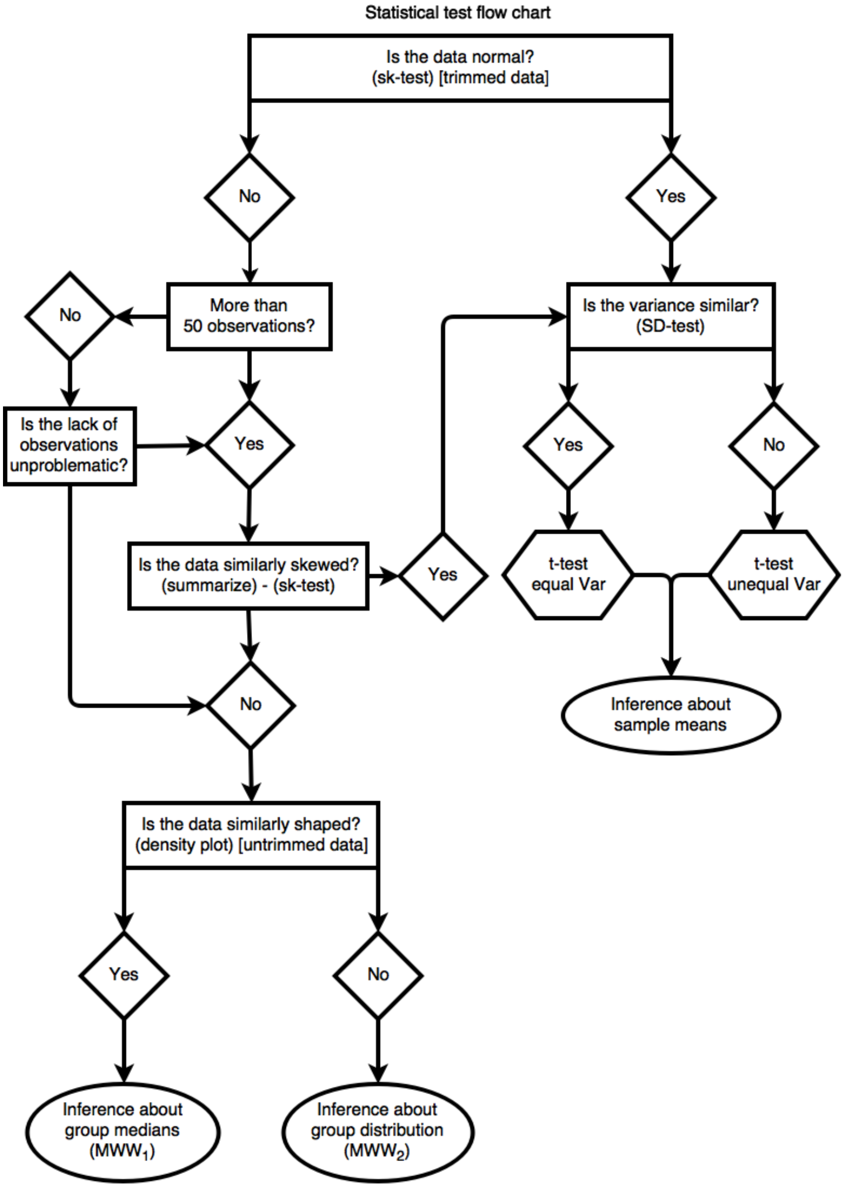


Figure 4 - Statistical Test Selection Flow Chart

8 Findings

In this section, we look at the effects of the different fund characteristics on IRR, size and TVPI. The main characteristic is fund type, and sequence number, cycles, decades and size will be controlled for both individually, and in combination with the main characteristic. Because of relatively few observations concerning quartile data, at least when we look back more than one period, quartile ranks and a discussion around them will be done separately.

8.1 IRR and TVPI

IRR and TVPI are two performance measures that will not show very different results given that the underlying characteristics are similar. We choose to use IRR when presenting this analysis and then comment if TVPI displays a different result.

8.1.1 Sequence Numbers

We see from Table XXX that all other sequence numbers show strong[6], significant differences from sequence number 0. Hence, we can say that sequence number 0 has a significantly lower distribution than all other funds. Consequently, it is adjacent to believe that GPs with sub-par performance close operations either straight away, or because they do not get the funding necessary to raise follow-on funds.

Sequence number 1 (12.47%) has the highest return of all sequence numbers. The results are weakly significant compared to sequence number 2 and 4. This is not surprising, as it is mostly GPs with the best performing first funds that raises a follow-on fund. Looking at TVPI, the result is strongly significant compared to fund number 4.

We find somewhat different results than Kaplan & Schoar (2005) looking at first-time funds (10.71%). While they find that first-time funds perform worse than funds with higher sequence number, our results point towards the opposite. Sequence number 0&1 has a higher performance compared to sequence number 2, 3 and 4, however, none of these results are

[6] 0.01 ≥ p-value = strongly significant. 0.01 < p-value ≤ 0.05 = significant. 0.05 < p-value ≤ 0.1 = weakly significant.

significant. This is surprising, as we would expect that when sequence numbers increase, sup-performing GPs would be weeded out, and thus leave us with a larger share of skilled GPs.

Comparing sequence number 2 (8.70%), 3 (10.16%) and 4 (9.46%), we find no significant differences between any of them. All results for sequence numbers seem to suggest that experience is not necessarily a contributing factor for good results. Kaplan & Schoar (2005) find that "*GPs of higher sequence number funds are better able to survive the poor performance of one particular fund*(source)", which may help us explain the somewhat surprising results.

8.1.2 Fund Characteristics

8.1.2.1 Fund Types

We find that buyout funds clearly outperform venture funds, averaging an IRR of 13.32% and 7.27%, respectively (Table 12 and 14). The result is strongly significant and coincides with the findings of Ljungqvist & Richardson (2003), Robinson & Sensoy (2011) and Harris, Jenkinson, & Kaplan (2014). As the two latter papers also have similar sample periods, the result is not surprising. Some of the other papers find that venture outperforms buyout. However, they do have little data after 2000, and as we will see later on, venture funds have not performed well during the 00s.

8.1.2.2 Cycles

As we hypothesised, funds raised in bust periods have a higher IRR (12.76%) and outperform funds raised in boom periods (9.18%). The result is significant, and in line with the finding of Kaplan and Schoar (2005). They find evidence that funds raised in periods of high economic growth are less likely to create follow-on funds. In our opinion, this implies lower performance for funds raised in boom periods. Barber & Yasuda (2010) find that fundraising often start when interim performances of GPs' existing funds are peaking. We believe these peaks usually happens during boom periods, thus implying our results are similar.

Comparing cycles on buyout, the median IRRs of funds raised in boom and bust periods, are 11.10% and 17.00%, respectively. This result is strongly significant, but looking at TVPI, the result is not significant at all, leading us to assume that the duration of funds raised in a bust period is smaller than for funds raised in a boom period. Without access to cash flow data, we have no way of confirming or rejecting this assumption. Venture funds raised in boom periods average an IRR of 6.71% compared to those raised in bust periods that average 10.03%. The difference is weakly significant.

Comparing fund types on cycles, buyout clearly outperforms venture in boom periods. This is strongly significant. In bust periods, however, we can only tell that the distribution of buyout performance is larger than that of venture. This result, too, is strongly significant. A closer look at IRR and TVPI for these funds, all suggest that buyout funds outperform venture funds when raised in bust periods.

8.1.2.3 Decades

We find that funds raised in the 90s clearly outperform funds raised in the 00s. Having an IRR of 12.11% and 7.87%, respectively, the result is strongly significant. Higson & Stucke (2012) find a downward trend in performance looking at vintage returns. Even though this seem to coincide with our result, they only study buyout funds.

When we look at buyout funds and compare decades, we find higher performance in the 00s (14.07%) than in the 90s (12.52%). Although the result is not significant, it does not show a downward trend in buyout performance over the years. This coincides with Harris, Jenkinson & Kaplan’s (2014) study. In addition to finding high performance for buyout funds throughout both decades, they also report significant differences in venture fund performance. We find an average IRR of 11.78% for venture funds raised in the 90s, and an IRR for funds raised in the 00s at 4.20%. The difference is strongly significant, and as mentioned earlier, we believe this is mainly due to the effects of the dot-com bubble.

Looking at different cycles and comparing fund types, we see that there is no significant difference between buyout and venture in the 90s. However, the difference in the 00s is strongly significant, which is expected, given the results mentioned above.

8.1.2.4 Size

Last, we compare sizes. Small, medium and large funds have an IRR of 10.68%, 9.85% and 8.82%, respectively. None of these performance metrics are significantly different from each other, but looking at TVPI, medium funds perform better than large funds. This difference is weakly significant. These results do not match those of Kaplan & Schoar (2005). They find that larger funds perform better than small. Robinson & Sensoy (2011) also find a somewhat different relationship between performance and size, namely that it is concave.

However, both Ljungqvist & Richardson (2003) and Gompers & Lerner (2000) find results suggesting higher performance for smaller funds. Although our results are not significant, we have similar findings. The results of the latter study indicates that larger funds pay higher prices for their investments, thus performing worse than smaller funds.

Taking a closer look at buyout funds, we can see that there are no significant differences in performance for any fund sizes. Looking at TVPI, however, we can see that small buyout funds have a larger distribution than both medium and large funds, while medium funds outperform large funds. Our results clearly contradicts the suggestion of Higson & Stucke (2012), who suggest that large buyout funds perform better than small, due to easier access to debt financing. Metrick & Yasuda's (2010) findings that GPs of buyout funds with some experience favour quantity over quality, seem to be a plausible explanation.

For venture funds, there are no significant differences when comparing small and medium funds. However, our results show that small funds have a significantly larger distribution than large funds. Medium venture funds also outperform large venture funds. These results contradicts those of Harris, Jenkinson & Kaplan (2014), who find that smaller venture funds underperform compared to larger venture funds. Comparing fund types, we can say that small buyout funds have a larger distribution than small venture funds with medians of 17.00% and 6.80%, respectively. This is weakly significant, but for TVPI, the result is strongly significant. Looking at medium and large funds, buyout clearly outperform venture funds. These results are also strongly significant.

8.2 Size

8.2.1 Sequence Numbers

Our results show that sequence number 0 has a significant lower size distribution than any other sequence numbers, except for sequence number 1, where there is no significant difference.

We see that sequence number 1 ($294M) is significantly lower than both sequence number 3 ($401M) and 4 ($526M). These results are not surprising, as we expected funds with higher sequence numbers to be larger. Looking at sequence number 0&1 ($309M), we find almost identical results as for sequence number 1. The only difference is that, compared to sequence number 3, the difference in size is now only weakly significant.

Comparing sequence number 2 ($337M) and 3, we find no significant differences. However, both these funds are significantly smaller than sequence number 4. The results show that sequence number 4 is significantly larger than all other sequence numbers. Knowing that sequence number 4 also contain funds with higher sequence numbers, this is not surprising. Metrick & Yasuda's (2010) finding, that GPs in buyout funds with experience sharply increase the size of their funds, is in line with our results.

8.2.2 Fund Characteristics

8.2.2.1 Fund Types

Looking at fund types, buyout ($545M) is significantly larger than venture ($306M). Again, the finding mentioned in the former paragraph by Metrick & Yasuda (2010), supports this. Higson & Stucke's (2012) suggestion that buyout funds are more scalable, also backs this result. In addition, we believe that venture funds benefit from being smaller. Hence, the result is not surprising.

8.2.2.2 Cycles

In our hypothesis, we believe that funds raised in boom periods would be larger than funds raised in bust period due to cautious investors. The results show that boom ($406M) is only fractionally larger than bust ($386M), however, this is not significant. In retrospect, we may have based our hypothesis on investors' behaviour in a too generalised way. Institutional investors might be less cautious than non-professionals about investing in bust periods. They have longer investment horizons and may see a market correction as an opportunity. A longer horizon may also enable them to sit through economic downturns.

Our results also show that buyout funds raised in boom periods are larger than buyout funds raised in bust periods. With a median of $456M and $269M, respectively, the difference is weakly significant. Looking at venture funds, we find the opposite, although the result is not significant. Comparing fund types, we can see that buyout funds are significantly larger than venture in boom periods. Looking at bust periods, we can only say that the distribution of buyout funds is significantly larger than that of venture funds. This indicates that more investors turn to venture funds in bust periods. If investors, too, believe that venture returns are more cyclical, there is a huge potential upside to these investments when prices are low.

8.2.2.3 Decades

A closer look at decades tells us that funds raised in the 00s ($413M) are a fraction larger than funds raised in the 90s ($390M). This result is not significant and does not coincide completely with the findings of Harris, Jenkinson & Kaplan (2014). They find that on average, fund sizes increase independent of fund types. Considering the growth of the PE industry over the last decades, our results implies that there are a lot more funds in the 00s. We wrote in our hypothesis that this result was possible, and argued that two bust periods in the 00s could be the reason why. However, looking at the results concerning size and cycles, this does not seem to be the case. Another possible explanation may be that GPs do not want to invest in too many companies, as it would be harder to be equally dedicated to all of them. Hence, funds do not need to be any larger than earlier.

Comparing decades, we see that buyout funds are somewhat larger in the 90s. It does not seem like the buyout boom in the 00s have had that big of an impact on the size of buyout funds. We find that the average decrease for buyout funds from the 90s to the 00s has been 10.5%, but using Harris, Jenkinson & Kaplan’s (2014) numbers we find an increase of 81.6%. The difference is huge, but we do not know whether their numbers are adjusted for inflation, or if they have trimmed the dataset for outliers.

What we find more surprising, though, is that venture funds raised in the 90s ($239M) are significantly smaller than venture funds raised in the 00s ($352M). Looking at it in retrospect, we see that funds raised in 2000 and in the beginning of 2001, just before the bubble burst, may have helped increasing the average of fund sizes in the 00s. Again, looking at Harris, Jenkinson & Kaplan’s (2014) study, their numbers suggest an increase in mean, which is much bigger than what we find (87.4% vs. 47.2%).

Comparing fund types on decades, we find that buyout funds have a larger distribution than venture in the 90s. In the 00s, buyout is clearly larger than venture. This result is strongly significant.

8.3 Quartile Data

There is no statistical difference between a current fund's performance if a GP's last fund was in the first or second quartile. However, there is a weak significant difference if a GP's last fund was in the top half versus the bottom half. We also see that if a GP's last fund was in the bottom quartile, their current fund perform significantly worse compared to the current funds of those GPs who's last fund was in any of the three higher quartiles.

Looking at all buyout and venture funds, independent of GP location and region focus, quartile results are consistent with those findings we got from US buyout and venture funds. This is also true if we look at all funds regardless of GP location and region focus.

This leads us to believe that picking top performers based on a GP's last fund, is not possible. However, the performance of a GP's last fund can be used as an indicator of which funds to avoid.

To look for stronger persistence, we need to include data from all buyout and venture funds regardless of GP location and region focus, in order to increase number of observations. In the two-period case, we compare current fund performance of GPs with two previous consecutive top quartile funds, against current fund performance of GPs that that did not have a top quartile fund two periods ago. In the three-period case, we compare current fund performance of GPs with three previous consecutive top quartile funds, against current fund performance of GPs that that did not have a top quartile fund two or three periods ago.

In the two-period case, the consecutive top performers perform statistically better than the current fund of GPs with their last fund in the second quartile or below. In the three-period case, the current fund of consecutive top performing GPs delivers significantly better results than the current fund of GPs that only have their last fund in the top quartile, but none of their previous funds ranked top quartile. They also perform better than GPs with their last fund in the second, third or fourth quartile.

However, this is only true if we look at IRR. If TVPI is used as the performance metric, the three-period case is less significant than the two-period case. In addition, we cannot claim that the current fund of consecutive top performers is statistically different from the current performance of GPs with their last fund in the first or second quartile.

This is a little surprising, but has two possible reasons. Either there are too few observations, or, because the IRR is the main factor in the construction of quartile ranks, it might make it unsuitable for inferences about TVPI persistence.

The results seem to be in line with the conclusions of Hendershott's (2008) study. If there is roughly a 40% probability that a fund in the top quartile is being followed by another top quartile fund, this is above the 25% of funds that would have been able to follow a top quartile fund if it followed a random selection. Hendershott argues that this is because there is a higher amount of exceptionally able GPs present in the top quartile. We therefore expect the probability, that a GP with two or more consecutive top quartile funds raise a new top quartile funds, to be more than 40%.

Looking deeper into this, we restrict our data to GPs with more than one sequence number and corresponding quartile rank. We look at the chances that the next fund is top quartile given that a GP has had consecutive top quartile funds. As the number of past top quartile funds increases, we expect the chance to increase. We also look at the chances of beating the benchmark, achieving an IRR above the top quartile return of all observations in our data set (20.1%) and the chance of achieving a positive return.

The analysis show that chances for a GP's next fund to be ranked in the top quartile increase with the number of consecutive top quartile performances. The chances are 35%, 41% and 59% (See Appendix H), depending on how many top performing funds a GP has had. These results are similar to those found by Hendershott (2008). We also see that as the number of consecutive top performances increase, so does the likelihood of beating the benchmark IRR and the chance of achieving a positive return.

8.3.1 General Note on Persistence

A potential problem with linking past performance to future performance is that it does not take in to account the risk profiles of GPs. A venture fund focusing on new technology or early stage pharmaceutical companies might on average deliver greater returns than venture funds focusing on retail or distribution. The former is potentially more risky and thus increases the risk of a fund, run by GPs with superior abilities, falling outside the top quartile rank.

The same is true for buyout. If some buyout GPs consistently use more leverage than others with similar abilities, the former strategy may on average generate higher returns, however, it is also more risky. This leads us to assume that a choice of consecutive top performers in a PE investment portfolio can be used as a way of reducing risk, more so than an effective way of picking top performers.

8.4 Data points outside our sample

After running our tests, and considering previous studies on the field, we ran some additional tests to see if the inferences from our study could be generalised to the whole PE investment universe. We focused on whether there was a significant difference between performance for funds focusing on investing in the US and funds with other countries or regions as their main focus area. Later we compared results from US buyout and venture funds to the other fund types also covered in the original dataset.

8.4.1 Regions

By focusing on different regions, we classified the different focus areas into continents in order to increase observations. Due too generally few observations and non-normal data, we ran the MWW-test to see if there were any differences in distributions. Of the six continents we examined, only Oceania and South America had significantly different distributions than the US. We can see from this that funds with a focus in these two continents deliver significantly better results than funds focusing on the US market alone.

Looking closer at Oceania, we see an overweight of small funds (58%) compared to the US [22%]. We also see an overweight of buyout funds (38%) [21%] and funds with sequence number 1 (31%) [21%]. Oceanian venture funds make up 27% [33%] of the total fund types. Since we have found that Oceania have more funds with characteristics associated with high performance, it is not surprising that funds focusing on this continent outperform the US.

If we look closer at South America there is significant differences if we use TVPI as our performance metric. If IRR is used it is not significantly different from the US. We would expect this insignificance to be due to the lack of observations, but we cannot be certain that this is the case.

From Table we see that the fund size dispersion is fairly similar to the US, only with a slightly higher tendency towards small funds (32%) [22%]. We find a substantial overweight of funds started in the 00s looking at South American funds (89%) [61%]. This should have led to a lower performance, since the 90s outperform the 00s in the main analysis. However, the same macroeconomic factors may not apply to this region.

Looking at fund types, there is a low representation of venture funds, but a substantial overweight of infrastructure (21%) [1%] and growth funds (37%) [2%] focusing on this region. We find no evidence that these funds deliver different returns compared to buyout and venture. However, there is a possibility that the economic conditions in this region makes these fund types more suitable. Looking deeper into the data, we find that most growth funds underperform compared to the South American median (1.74) with five out of seven funds being below the median. If the US median is used [1.45] three out of seven are still below. However, all infrastructure and buyout funds perform better than the US median and all infrastructure funds also perform better than the US median. This leads us to believe that infrastructure and buyout is highly suited for this area, but growth funds cannot help us explain the difference. A closer description of continent fund composition is found in Appendix G.

8.4.2 Fund Types

After going through all fund types, we only find that distressed debt and natural resources deliver significantly better results than buyout and venture funds. This is only considering IRR, though. If we use a one-sided MWW-test and look at TVPI, natural resources are not statistically different from US buyout and venture funds. The two-tailed test is inconclusive with a p-value of 1.2484. We find the same with distressed debt, which returns a p-value of 1.8525. These probabilities should not be possible, but with the way Microsoft Excel, and most statistical software calculates two-tailed tests, this might happen if we have non-symmetric distributions (Kulinskaya, 2008).

If we look more closely at these funds, we find that 75% of US distressed debt funds have an IRR above the combined buyout and venture median IRR. The equivalent measure for natural resources is 85%. Since TVPI is not significantly different for any of these, it leads us to believe that the average lifetime of these funds are shorter than those of buyout and venture funds. Looking at the cash flow charts available for the funds in question, it does not seem like this is the case. However, without full access to all cash flow data, we have no way of confirming or rejecting this hypothesis. A closer dicription of fund comparison can be found in Appendix G

Based on the findings from the data outside our sample, it seems like most of the findings for US based buyout and venture, can be generalised to other fund types and other regions. However, there is a problem with lack of observations, and a quantitative analysis will encounter problems. An in-depth case study might be a better way of researching these less explored fund types and regions.

9 Conclusion

In this thesis, we have investigated the performance of private equity funds, focusing on the buyout and venture sector. We have used a dataset of individual fund returns and characteristics form the Preqin database over the period from 1990 to 2008. Most other private equity papers have focused on funds raised up until the late 1990s and early 2000s, but the recent inflow of capital to this sector has created a need for a re-evaluation of fund and GP performance.

We find that experience and past performance is not necessarily the best determinant for future fund performance. It is, however, a likely determinant of future fund size. Based on the growth of follow-on funds, the ability of GPs with a higher sequence number to survive a poor performing fund, leads to the conclusion that too much weight is being put on past performance. Even though our data does not show strong signs of PE persistency, picking GPs with strong past performance could be used in a risk reduction strategy.

Secondly, we find that buyout funds outperform venture funds on a general basis. Actually, we find no significant results showing that venture outperform buyout no matter what we control for. We believe that this is much due to the buyout boom in the 00s and the fact that debt financing has been a lot cheaper in later years.

Thirdly, there are few differences in our findings when comparing IRR results to TVPI. This could indicate that there are no systematic underlying factors, or timing differences, that affect the two fund types. Hence, the IRR is, on average, a reliable performance metric.

Fourthly, we find that our results can be generalised to most parts of the PE investment universe. Most regions perform similar to the US, and if there is a discrepancy, it is mostly because of the difference in fund composition. There might be some differences due to legislation, but this is outside the scope of our thesis and will need further research. We also find that most funds, on average, emulate the return characteristics of buyout and venture combined. The only discrepancies here are natural resources and distressed debt. We cannot find any particular reason for this due to lack of data.

We find that our results using the Preqin dataset show similar results to more recent studies on private equity, especially those using the Burgiss dataset. Few papers have focused specifically on sequence numbers. The paper from Kaplan & Schoar (2005) is one of the few that have focused on this, but the study is now ten years old and the VE dataset has been shown to exhibit some weaknesses. Our thesis includes some results on sequence numbers, but a more in-depth study of sequence numbers could highlight some important characteristic of specific GPs behaviour, enabling LPs to pick better PE funds in the future. As the Preqin dataset lends itself to easy sequence number calculations, this would be a great way of expanding the knowledge of PE performance.

10 References

Acharya, V. V, Franks, J., & Servaes, H. (2007). Private equity: Boom and bust? *Journal of Applied Corporate Finance*, *19*(4), 44–53.

Appelbaum, E., & Batt, R. (2012). A Primer on Private Equity at Work. *Challenge*, *55*(5), 5–38.

Barber, B. M., & Yasuda, A. (2014). Interim Fund Performance and Fundraising in Private Equity. Retrieved from Available at SSRN 2357570

Bureau of Labor statistics. (2015). CPI Inflation Calculator. Retrieved March 7, 2015, from http://www.bls.gov/data/inflation_calculator.htm

Cendrowski, H., Petro, L. W., Martin, J. P., & Wadecki, A. A. (2012). *Private equity: History, governance, and operations* (Second edi.). Hoboken, New Jersey: John Wiley & Sons, Inc.

CFA Institute. (2010). *Global Investment Performance Standard*. Retrieved from http://www.cfapubs.org/doi/pdf/10.2469/ccb.v2010.n5.1

Chung, J.-W., Sensoy, B. A., Stern, L., & Weisbach, M. S. (2012). Pay for performance from future fund flows: the case of private equity. *Review of Financial Studies*.

Gilligan, J., & Wright, M. (2014). *PRIVATE EQUITY DEMYSTIFIED An explanatory guide Second edition* (Third.). Retrieved from http://www.icaew.com/~/media/corporate/files/technical/corporate finance/financing change/tecplm12976_privateequityiii_full edition.ashx

Gompers, P., & Lerner, J. (2000). Money chasing deals? The impact of fund inflows on private equity valuation. *Journal of Financial Economics*, *55*(2), 281–325. doi:10.1016/S0304-405X(99)00052-5

Gupta, V. (2012). Benchmarking Private Equity Getting Through the Maze. *Russell Research*.

Hall, D. (2006). Private equity and infrastructure funds in public services and utilities. *PSIRU Reports*.

Harris, R., Jenkinson, T., & Kaplan, S. N. (2014). Private equity performance: What do we know? *The Journal of Finance*, *69*(5), 1851–1882.

Harris, R., Jenkinson, T., Kaplan, S., & Stucke, R. (2014). Has persistence persisted in private equity? Evidence from buyout and venture capital funds. *Working Paper, University of Chicago*.

Harris, R., Jenkinson, T., & Stucke, R. (2010). A White Paper on Private Equity Data and Research. *UAI Foundation Consortium Working Paper, Univerity of Virginia*, (December).

Hellmann, T., & Puri, M. (2002). Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence. *The Journal of Finance*, *57*(1), 169–197. doi:10.1111/1540-6261.00419

Hendershott, R. (2008). *Using Past Performance to Infer Investment Manager Ability*. *Working Paper, Leavey School of Business, Santa Clara University*.

Higson, C., & Stucke, R. (2012). The performance of private equity. *Working Paper, Collier Institute of Private Equity, London Business School*.

Hsu, D. H. (2004). What do entrepreneurs pay for venture capital affiliation? *The Journal of Finance*, *59*(4), 1805–1844.

Kaplan, S. N., & Schoar, A. (2005). Private Equity Performance: Returns, Persistence, and Capital Flows. *The Journal of Finance*, *60*(4), 1791–1823.

Kulinskaya, E. (2008). On two-sided p-values for non-symmetric distributions. *arXiv Preprint arXiv:0810.2124*.

Ljungqvist, A., & Richardson, M. (2003). *The cash flow, return and risk characteristics of private equity* (No. 9454). Retrieved from http://www.nber.org/papers/w9454

Metrick, A., & Yasuda, A. (2010). The economics of private equity funds. *Review of Financial Studies*, *23*, 2303–2341.

Metrick, A., & Yasuda, A. (2011). *Venture Capital & the Finance of Innovation*. (L. Vitetta, Ed.) (Second Edi.). John Wiley & Sons, Inc.

the National Bureau of Economic Research. (2015). US Business Cycle Expansions and Contractions. Retrieved March 12, 2015, from http://www.nber.org/cycles/cyclesmain.html

Phalippou, L. (2008). The Hazards of Using IRR to Measure Performance: The Case of Private Equity. *Journal of Performance Measurement*, *12*(4), 55–67.

Phalippou, L., & Gottschalg, O. (2008). The Performance of Private Equity Funds. *Review of Financial Studies*.

Preqin. (2008). March 2008 PE Spotlight. *Private Equity Spotlight*, *4*(3). Retrieved from https://www.preqin.com/docs/newsletters/pe/SpotlightMarch2008.pdf

Preqin. (2014). *Preqin AnnouncestheMost Consistent Performing Private Equity Fund Managers in 2014*. Retrieved from https://www.preqin.com/docs/press/Consistent-PE-14.pdf

Preqin. (2015). *Preqin Industry News*. Retrieved from https://www.preqin.com/docs/newsletters/pe/Preqin-PESL-May-15-Industry-News.pdf

Robinson, D. T., & Sensoy, B. A. (2011). *Cyclicality, Performance Measurement, and Cash Flow Liquidity in Private Equity* (No. 17428). Retrieved from http://www.nber.org/papers/w17428

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *The Journal of Finance*, *53*(5), 1589–1622. Retrieved from http://www.jstor.org/stable/117418

Steer, J., & Ellis, C. (2011). Are UK venture capital and private equity valuations over-optimistic? *BVCA Research Report*.

Steven N. Kaplan, & Strömberg, P. (2008). Leveraged Buyouts and Private Equity. *Journal of Economic Perspective*, *22*(4).

Stucke, R. (2011). Updating history. *Working Paper*, *Oxford University*.

Wooldridge, J. M. (2014). *Introduction to Econometrics*. (A. Ainscow, A. Burt, & F. Rowe, Eds.) (EMEA.). Andover: Cengage Learning EMEA.

Appendices

Appendix A – Descriptive Statistics for Trimmed Dataset

Descriptive Statistics for IRR											
Sequence Number	Statistics	All	Buyout	Venture	Boom	Bust	90s	00s	Small	Medium	Large
All	Observations	668	271	397	567	101	292	376	127	360	181
	Mean	9.73	13.32	7.27	9.18	12.76	12.11	7.87	10.68	9.85	8.82
	Median	8.65	12.00	5.00	8.20	11.40	9.65	7.85	8.90	8.10	8.90
	Std Error	0.55	0.72	0.75	0.60	1.35	0.96	0.61	1.40	0.77	0.87
0&1	Observations	177	89	88	155	22	95	82	57	87	33
	Mean	10.71	12.66	8.74	9.86	16.69	11.90	9.33	12.67	9.22	11.24
	Median	10.00	11.50	8.70	9.70	14.00	9.70	10.05	12.00	9.30	10.70
	Std Error	1.01	1.27	1.55	1.04	3.32	1.54	1.24	2.08	1.28	2.22
0	Observations	39	14	25	33	6	23	16	15	16	8
	Mean	4.46	8.64	2.13	4.03	6.87	5.44	3.06	7.33	3.33	1.36
	Median	2.90	5.90	1.40	2.20	4.00	3.90	0.45	7.10	2.40	-2.10
	Std Error	2.07	4.02	2.24	2.36	3.70	2.96	2.78	4.21	2.58	3.75
1	Observations	138	75	63	122	16	72	66	42	71	25
	Mean	12.47	13.41	11.36	11.44	20.37	13.97	10.84	14.58	10.55	14.40
	Median	10.70	11.70	9.70	10.55	15.25	11.60	10.60	12.40	10.00	12.10
	Std Error	2.09	3.02	2.99	2.15	6.57	3.08	1.33	2.35	1.41	2.38
2	Observations	159	72	87	138	21	72	87	37	90	32
	Mean	8.70	12.25	5.76	7.95	13.61	8.91	8.53	8.66	9.61	6.17
	Median	7.70	11.65	3.50	7.50	11.40	7.00	8.20	7.00	9.50	7.25
	Std Error	2.04	3.03	2.62	2.18	5.41	3.11	1.14	2.17	1.45	1.47
3	Observations	110	47	63	85	25	47	63	17	66	27
	Mean	10.16	16.11	5.72	10.22	9.96	12.53	8.39	8.41	9.18	13.64
	Median	8.90	15.40	3.90	8.90	9.40	9.20	8.90	8.90	7.15	14.40
	Std Error	2.24	2.60	3.02	2.59	4.40	3.27	1.57	4.21	1.78	1.90
4	Observations	222	63	159	189	33	78	144	16	117	89
	Mean	9.46	13.40	7.90	9.07	11.73	15.07	6.43	10.63	10.87	7.41
	Median	7.50	12.10	4.80	6.40	11.60	10.80	5.95	3.60	7.70	8.20
	Std Error	2.51	3.61	3.25	2.78	5.74	4.03	1.03	4.99	1.59	1.30

Table 3 - Descriptive Statistics for IRR - Trimmed Dataset - All mean and median numbers are in percentages

Descriptive Statistics for Size											
Sequence Number	Statistics	All	Buyout	Venture	Boom	Bust	90s	00s	Small	Medium	Large
All	Observations	668	271	397	567	101	292	376	127	360	181
	Mean	403	545	306	406	386	390	413	58	259	931
	Median	266	402	204	271	252	239	277	57	245	803
	Std Error	15	28	16	17	39	24	20	2	6	30
0&1	Observations	177	89	88	155	22	95	82	57	87	33
	Mean	309	448	169	316	259	283	340	56	215	995
	Median	171	264	108	178	92	162	172	57	194	844
	Std Error	29	50	23	32	77	35	49	3	10	78
0	Observations	39	14	25	33	6	23	16	15	16	8
	Mean	363	675	188	377	285	263	507	54	185	1298
	Median	119	375	108	120	78	104	120	52	137	1339
	Std Error	85	182	65	95	190	76	173	6	27	167
1	Observations	138	75	63	122	16	72	66	42	71	25
	Mean	294	405	162	300	249	289	299	57	222	897
	Median	182	264	108	191	97	191	174	57	206	756
	Std Error	29	48	20	31	83	40	43	4	10	81
2	Observations	159	72	87	138	21	72	87	37	90	32
	Mean	337	484	216	347	272	358	320	61	266	859
	Median	245	395	163	245	252	231	267	58	249	725
	Std Error	26	45	23	29	48	43	32	3	12	64
3	Observations	110	47	63	85	25	47	63	17	66	27
	Mean	401	612	243	415	353	410	394	57	277	921
	Median	289	486	170	299	243	288	290	55	254	779
	Std Error	37	59	36	45	56	62	45	6	15	84
4	Observations	222	63	159	189	33	78	144	16	117	89
	Mean	526	701	456	518	570	539	519	65	276	937
	Median	383	570	335	380	435	367	407	62	262	858
	Std Error	28	62	29	30	87	54	33	6	10	40

Table 4 - Descriptive Statistics for Size – Trimmed Dataset - All mean and median numbers are in million USD

Descriptive Statistics for TVPI											
Sequence Number	Statistics	All	Buyout	Venture	Boom	Bust	90s	00s	Small	Medium	Large
All	Observations	668	271	397	567	101	292	376	127	360	181
	Mean	1.56	1.77	1.42	1.53	1.71	1.65	1.49	1.59	1.59	1.47
	Median	1.46	1.68	1.27	1.45	1.53	1.50	1.41	1.47	1.45	1.48
	Std Error	0.03	0.04	0.04	0.03	0.08	0.05	0.04	0.07	0.04	0.04
0&1	Observations	177	89	88	155	22	95	82	57	87	33
	Mean	1.62	1.75	1.49	1.57	2.00	1.64	1.60	1.72	1.57	1.59
	Median	1.58	1.71	1.45	1.52	1.83	1.58	1.57	1.72	1.47	1.53
	Std Error	0.06	0.07	0.08	0.06	0.20	0.08	0.08	0.11	0.07	0.12
0	Observations	39	14	25	33	6	23	16	15	16	8
	Mean	1.24	1.40	1.14	1.22	1.30	1.30	1.14	1.36	1.20	1.08
	Median	1.12	1.43	1.05	1.10	1.14	1.16	1.01	1.34	1.11	0.89
	Std Error	0.09	0.15	0.12	0.10	0.21	0.12	0.14	0.17	0.14	0.18
1	Observations	138	75	63	122	16	72	66	42	71	25
	Mean	1.73	1.82	1.63	1.66	2.26	1.75	1.71	1.84	1.66	1.75
	Median	1.64	1.77	1.59	1.60	2.08	1.71	1.63	1.77	1.63	1.64
	Std Error	0.06	0.08	0.10	0.06	0.24	0.09	0.09	0.14	0.08	0.13
2	Observations	159	72	87	138	21	72	87	37	90	32
	Mean	1.53	1.74	1.36	1.48	1.88	1.51	1.55	1.49	1.59	1.41
	Median	1.46	1.67	1.22	1.44	1.75	1.45	1.51	1.35	1.52	1.43
	Std Error	0.06	0.08	0.08	0.06	0.19	0.10	0.07	0.12	0.08	0.08
3	Observations	110	47	63	85	25	47	63	17	66	27
	Mean	1.62	1.90	1.40	1.63	1.57	1.74	1.52	1.54	1.60	1.70
	Median	1.44	1.88	1.17	1.48	1.38	1.54	1.39	1.32	1.39	1.69
	Std Error	0.08	0.09	0.12	0.10	0.14	0.14	0.09	0.24	0.12	0.10
4	Observations	222	63	159	189	33	78	144	16	117	89
	Mean	1.50	1.73	1.41	1.50	1.53	1.73	1.38	1.47	1.59	1.39
	Median	1.38	1.54	1.28	1.36	1.46	1.51	1.34	1.24	1.40	1.34
	Std Error	0.05	0.09	0.06	0.06	0.11	0.11	0.06	0.21	0.08	0.07

Table 5 - Descriptive Statistics for TVPI - Trimmed Dataset

Appendix B – Descriptive Statistics for Untrimmed Dataset

Descriptive Statistics for IRR											
Sequence Number	Statistics	All	Buyout	Venture	Boom	Bust	90s	00s	Small	Medium	Large
All	Observations	786	305	481	671	115	364	422	176	398	212
	Median	8.60	11.60	5.10	8.30	11.30	9.90	7.70	7.75	8.60	8.85
	Mean	12.80	12.55	12.96	12.91	12.17	19.62	6.91	11.72	15.75	8.16
	Std Error	1.20	0.88	1.88	1.37	1.64	2.39	0.75	2.21	2.08	0.96
0&1	Observations	208	94	114	183	25	113	95	80	94	34
	Median	9.70	11.25	7.70	9.60	11.60	9.70	9.70	9.70	9.15	11.10
	Mean	10.80	11.36	10.34	10.36	14.04	14.05	6.94	10.60	10.67	11.63
	Std Error	1.84	1.58	3.11	2.04	3.40	2.90	2.04	3.24	2.92	2.19
0	Observations	51	16	35	43	8	29	22	25	18	8
	Median	0.00	4.40	-2.20	-1.70	4.00	2.20	-3.55	1.40	1.10	-2.10
	Mean	-1.14	3.34	-3.18	-1.90	2.98	4.08	-8.00	-1.00	-2.44	1.36
	Std Error	3.98	5.25	5.29	4.64	4.73	5.35	5.75	7.42	4.57	3.75
1	Observations	157	78	79	140	17	84	73	55	76	26
	Median	11.60	11.90	10.10	10.70	15.00	11.60	10.70	13.00	10.05	12.15
	Mean	14.68	13.00	16.34	14.12	19.24	17.49	11.44	15.87	13.78	14.79
	Std Error	1.98	1.52	3.65	2.17	3.94	3.38	1.71	3.09	3.36	2.32
2	Observations	181	79	102	154	27	84	97	46	96	39
	Median	7.50	11.30	3.35	7.40	11.40	8.20	7.50	6.30	9.50	7.50
	Mean	11.55	12.25	11.01	11.23	13.40	17.34	6.54	11.13	13.38	7.55
	Std Error	2.35	1.70	3.97	2.69	3.66	4.76	1.35	4.46	3.74	2.66
3	Observations	128	51	77	102	26	59	69	25	71	32
	Median	8.75	14.40	3.90	8.75	9.00	9.20	8.60	8.90	7.20	11.55
	Mean	12.34	14.99	10.59	13.23	8.85	18.28	7.27	14.21	12.53	10.47
	Std Error	2.45	1.53	3.95	3.02	2.35	4.89	1.61	6.48	3.64	2.58
4	Observations	269	81	188	232	37	108	161	25	137	107
	Median	7.80	12.00	5.25	7.65	11.60	11.20	6.00	3.40	9.00	7.00
	Mean	15.40	12.69	16.57	15.89	12.34	27.96	6.98	13.87	22.57	6.58
	Std Error	2.51	1.98	3.50	2.87	3.37	5.85	1.12	5.45	4.66	1.26

Table 6 - Descriptive Statistics for IRR - Untrimmed Dataset - All mean and median numbers are in percentages

Descriptive Statistics for Size											
Sequence Number	Statistics	All	Buyout	Venture	Boom	Bust	90s	00s	Small	Medium	Large
All	Observations	786	305	481	671	115	364	422	176	398	212
	Median	252	447	178	257	226	227	273	56	236	884
	Mean	477	780	285	490	401	441	508	54	254	1249
	Std Error	26	59	16	30	48	34	39	2	5	74
0&1	Observations	208	94	114	183	25	113	95	80	94	34
	Median	140	262	83	143	82	134	147	54	188	871
	Mean	287	459	145	295	231	248	333	52	212	1050
	Std Error	29	54	19	31	69	31	51	3	9	94
0	Observations	51	16	35	43	8	29	22	25	18	8
	Median	104	273	83	111	58	83	114	57	140	1339
	Mean	295	606	153	309	222	223	391	54	184	1298
	Std Error	67	166	47	75	145	62	131	5	24	167
1	Observations	157	78	79	140	17	84	73	55	76	26
	Median	162	262	82	172	97	153	169	47	200	788
	Mean	285	429	142	291	235	257	316	50	218	973
	Std Error	31	56	17	34	79	36	53	4	10	109
2	Observations	181	79	102	154	27	84	97	46	96	39
	Median	224	415	146	229	224	221	245	57	245	779
	Mean	398	622	225	390	445	437	364	59	259	1141
	Std Error	42	81	33	42	150	65	55	4	11	139
3	Observations	128	51	77	102	26	59	69	25	71	32
	Median	257	542	164	276	228	258	257	53	249	799
	Mean	443	776	222	469	339	500	394	49	270	1134
	Std Error	53	110	31	65	56	99	49	5	14	153
4	Observations	269	81	188	232	37	108	161	25	137	107
	Median	371	764	294	367	415	325	441	57	260	963
	Mean	694	1309	429	720	529	613	748	54	270	1385
	Std Error	61	173	29	69	80	80	86	6	9	126

Table 7 - Descriptive Statistics for Size - Untrimmed Dataset - All mean and median numbers are in million USD

Descriptive Statistics for TVPI											
Sequence Number	Statistics	All	Buyout	Venture	Boom	Bust	90s	00s	Small	Medium	Large
All	Observations	786	305	481	671	115	364	422	176	398	212
	Median	1.46	1.66	1.28	1.46	1.51	1.52	1.40	1.41	1.46	1.48
	Mean	1.84	1.75	1.89	1.86	1.72	2.22	1.51	2.02	1.96	1.45
	Std Error	0.08	0.05	0.13	0.09	0.10	0.16	0.05	0.25	0.11	0.04
0&1	Observations	208	94	114	183	25	113	95	80	94	34
	Median	1.56	1.70	1.45	1.52	1.67	1.58	1.53	1.60	1.47	1.54
	Mean	1.75	1.75	1.75	1.73	1.86	1.76	1.73	1.82	1.73	1.62
	Std Error	0.10	0.09	0.17	0.11	0.20	0.13	0.16	0.18	0.16	0.12
0	Observations	51	16	35	43	8	29	22	25	18	8
	Median	1.00	1.23	0.90	0.90	1.14	1.10	0.83	1.05	1.05	0.89
	Mean	1.24	1.24	1.24	1.25	1.17	1.46	0.95	1.40	1.09	1.08
	Std Error	0.22	0.17	0.32	0.26	0.20	0.38	0.13	0.44	0.14	0.18
1	Observations	157	78	79	140	17	84	73	55	76	26
	Median	1.69	1.77	1.60	1.64	1.93	1.72	1.64	1.81	1.64	1.67
	Mean	1.91	1.85	1.97	1.88	2.19	1.86	1.97	2.02	1.88	1.78
	Std Error	0.11	0.09	0.20	0.12	0.23	0.12	0.20	0.17	0.19	0.13
2	Observations	181	79	102	154	27	84	97	46	96	39
	Median	1.46	1.65	1.19	1.42	1.65	1.48	1.40	1.32	1.52	1.45
	Mean	1.77	1.71	1.82	1.73	1.97	2.12	1.47	2.20	1.71	1.41
	Std Error	0.22	0.08	0.38	0.25	0.28	0.46	0.07	0.81	0.13	0.09
3	Observations	128	51	77	102	26	59	69	25	71	32
	Median	1.46	1.80	1.17	1.53	1.38	1.56	1.38	1.32	1.39	1.67
	Mean	1.92	1.86	1.95	2.02	1.53	2.41	1.49	2.23	1.94	1.61
	Std Error	0.20	0.09	0.33	0.25	0.14	0.41	0.10	0.56	0.30	0.10
4	Observations	269	81	188	232	37	108	161	25	137	107
	Median	1.40	1.55	1.33	1.39	1.46	1.55	1.35	1.21	1.46	1.34
	Mean	1.91	1.72	1.99	1.96	1.59	2.66	1.41	2.07	2.31	1.36
	Std Error	0.14	0.10	0.19	0.16	0.15	0.32	0.06	0.47	0.25	0.06

Table 8 - Descriptive Statistics for TVPI - Untrimmed Dataset

Appendix C – Sequence Number Comparisons

Sequence Number (SN) Comparisons by IRR				
SN	IRR (%)	Observations	P-value	Test
0 1	0.00 11.60	51 157	0.0000	MWW_2
0 2	0.00 7.50	51 181	0.0012	MWW_2
0 3	0.00 8.75	51 128	0.0005	MWW_2
0 4	0.00 7.80	51 269	0.0004	MWW_2
1 2	12.47 8.70	138 159	0.0123	T-test
1 3	12.47 10.16	138 110	0.1804	T-test
1 4	12.47 9.46	138 222	0.0502	T-test
0&1 2	10.71 8.70	177 159	0.1606	T-test
0&1 3	10.71 10.16	177 110	0.7392	T-test
0&1 4	10.71 9.46	177 222	0.3934	T-test
2 3	8.70 10.16	159 110	0.3758	T-test
2 4	8.70 9.46	159 222	0.5992	T-test
3 4	10.16 9.46	110 222	0.6951	T-test

Table 9 - Sequence Number Comparisons for IRR

Sequence Number (SN) Comparisons by Size				
SN	Size $M	Observations	P-value	Test
0 1	104 162	51 157	0.1198	MWW_2
0 2	104 224	51 181	0.0007	MWW_2
0 3	104 257	51 128	0.0001	MWW_2
0 4	104 371	51 269	0.0000	MWW_2
1 2	294 337	138 159	0.2700	T-test
1 3	294 401	138 110	0.0219	T-test
1 4	294 526	138 222	0.0000	T-test
0&1 2	309 337	177 159	0.4769	T-test
0&1 3	309 401	177 110	0.0532	T-test
0&1 4	309 526	177 222	0.0000	T-test
2 3	337 401	159 110	0.1586	T-test
2 4	337 526	159 222	0.0000	T-test
3 4	401 526	110 222	0.0097	T-test

Table 10 - Sequence Number Comparisons for Size

Sequence Number (SN) Comparisons by TVPI				
SN	TVPI	Observations	P-value	Test
0 1	1.00 1.69	51 157	0.0000	MWW_1
0 2	1.00 1.46	51 181	0.0006	MWW_2
0 3	1.00 1.46	51 128	0.0003	MWW_2
0 4	1.00 1.40	51 269	0.0005	MWW_2
1 2	1.73 1.53	138 159	0.0216	T-test
1 3	1.73 1.62	138 110	0.2662	T-test
1 4	1.73 1.50	138 222	0.0074	T-test
0&1 2	1.62 1.53	177 159	0.2650	T-test
0&1 3	1.62 1.62	177 110	0.9621	T-test
0&1 4	1.62 1.50	177 222	0.1270	T-test
2 3	1.53 1.62	159 110	0.3954	T-test
2 4	1.53 1.50	159 222	0.7130	T-test
3 4	1.62 1.50	110 222	0.2305	T-test

Table 11 - Sequence Number Comparisons for TVPI

Appendix D - Fund Characteristic Comparisons

Fund Characteristic (FC) Comparisons by IRR				
FC	IRR (%)	Observations	P-value	Test
Buyout Venture	13.32 7.27	271 397	0.0000	T-test
Boom Bust	9.18 12.76	567 101	0.0190	T-test
90s 00s	12.11 7.87	292 376	0.0002	T-test
Small Medium	10.68 9.85	127 360	0.5914	T-test
Small Large	10.68 8.82	127 181	0.2605	T-test
Medium Large	9.85 8.82	360 181	0.3768	T-test

Table 12 - Fund Characteristic Comparisons by IRR

Fund Characteristic (FC) Comparisons by Size				
FC	Size $M	Observations	P-value	Test
Buyout Venture	545 306	271 397	0.0000	T-test
Boom Bust	406 386	567 101	0.6478	T-test
90s 00s	390 413	292 376	0.4667	T-test

Table 13 - Fund Characteristic Comparisons by Size

Fund Characteristic (FC) Comparisons by TVPI				
FC	TVPI	Observations	P-value	Test
Buyout Venture	1.77 1.42	271 397	0.0000	T-test
Boom Bust	1.53 1.71	567 101	0.0297	T-test
90s 00s	1.65 1.49	292 376	0.0098	T-test
Small Medium	1.59 1.59	127 360	0.9600	T-test
Small Large	1.59 1.47	127 181	0.1659	T-test
Medium Large	1.59 1.47	360 181	0.0648	T-test

Table 14 - Fund Characteristic Comparisons by TVPI

Appendix E - Fund Type Comparisons on Fund Characteristics

Fund Type (FT) Comparisons on Fund Characteristics (FC) by IRR					
FC	FT	IRR (%)	Observations	P-value	Test
Boom	Buyout Venture	12.63 6.71	237 330	0.0000	T-test
Bust	Buyout Venture	17.00 6.65	37 78	0.0001	MWW_2
90s	Buyout Venture	12.52 11.78	131 161	0.6915	T-test
00s	Buyout Venture	14.07 4.20	140 236	0.0000	T-test
Small	Buyout Venture	17.00 6.80	25 151	0.0616	MWW_2
Medium	Buyout Venture	13.56 7.57	137 223	0.0001	T-test
Large	Buyout Venture	12.15 3.41	112 69	0.0000	T-test

Table 15 - Fund Type Comparisons on Fund Characteristics by IRR

Fund Type (FT) Comparisons on Fund Characteristics (FC) by Size					
FC	FT	Size $M	Observations	P-value	Test
Boom	Buyout Venture	558 267	237 330	0.0000	T-test
Bust	Buyout Venture	269 179	37 78	0.0168	MWW_2
90s	Buyout Venture	462 148	148 216	0.0000	MWW_2
00s	Buyout Venture	516 352	140 236	0.0001	T-test

Table 16 - Fund Type Comparisons on Fund Characteristics by Size

Fund Type (FT) Comparisons on Fund Characteristics (FC) by TVPI					
FC	FT	TVPI	Observations	P-value	Test
Boom	Buyout Venture	1.75 1.38	237 330	0.0000	T-test
Bust	Buyout Venture	1.77 1.36	37 78	0.0027	MWW_1
90s	Buyout Venture	1.73 1.58	131 161	0.1317	T-test
00s	Buyout Venture	1.80 1.30	140 236	0.0000	T-test
Small	Buyout Venture	2.15 1.32	25 151	0.0073	MWW_2
Medium	Buyout Venture	1.80 1.46	137 223	0.0001	T-test
Large	Buyout Venture	1.66 1.17	112 69	0.0000	T-test

Table 17 - Fund Type Comparisons on Fund Characteristics by TVPI

Appendix F - Fund Characteristic Comparisons on Fund Type

Fund Characteristic (FC) Comparisons on Buyout by IRR				
FC	IRR (%)	Observations	P-value	Test
Boom Bust	11.10 17.00	237 37	0.0028	MWW_1
90s 00s	12.52 14.07	131 140	0.2875	T-test
Small Medium	17.00 12.25	25 137	0.3732	MWW_2
Small Large	17.00 11.10	25 112	0.1608	MWW_2
Medium Large	13.56 12.15	137 112	0.3310	T-test

Table 18 - Fund Characteristic Comparisons on Buyout by IRR

Fund Characteristic (FC) Comparisons on Buyout by Size				
FC	Size $M	Observations	P-value	Test
Boom Bust	456 269	268 37	0.0551	MWW_1
90s 00s	576 516	131 140	0.2774	T-test

Table 19 - Fund Characteristic Comparisons on Buyout by Size

Fund Characteristic (FC) Comparisons on Buyout by TVPI				
FC	TVPI	Observations	P-value	Test
Boom Bust	1.63 1.77	268 37	0.1496	MWW_1
90s 00s	1.73 1.80	131 140	0.3891	T-test
Small Medium	2.15 1.69	25 142	0.0862	MWW_2
Small Large	2.15 1.58	25 138	0.0113	MWW_2
Medium Large	1.80 1.66	142 138	0.0998	T-test

Table 20 - Fund Characteristic Comparisons on Buyout by TVPI

Fund Characteristic (FC) Comparisons on Venture by IRR				
FC	IRR (%)	Observations	P-value	Test
Boom Bust	6.71 10.03	330 67	0.0994	T-test
90s 00s	11.78 4.20	161 236	0.0000	T-test
Small Medium	9.20 7.57	105 223	0.3850	T-test
Small Large	6.80 0.25	151 69	0.01914	MWW_2
Medium Large	7.57 3.41	223 69	0.0153	T-test

Table 21 - Fund Characteristic Comparisons on Venture by IRR

Fund Characteristic (FC) Comparisons on Venture by Size				
FC	Size $M	Observations	P-value	Test
Boom Bust	297 354	330 67	0.2687	T-test
90s 00s	239 352	161 236	0.0002	T-test

Table 22 - Fund Characteristic Comparisons on Venture by Size

Fund Characteristic (FC) Comparisons on Venture by TVPI				
FC	TVPI	Observations	P-value	Test
Boom Bust	1.38 1.61	330 67	0.0293	T-test
90s 00s	1.58 1.30	161 236	0.0013	T-test
Small Medium	1.48 1.46	105 223	0.8422	T-test
Small Large	1.32 1.01	151 74	0.0193	MWW_2
Medium Large	1.46 1.17	223 69	0.0005	T-test

Table 23 - Fund Characteristic Comparisons on Venture by TVPI

Appendix G – Data Outside our Sample

Data Outside Our Sample - Continent Comparison						
	US		Oceania		S America	
Type	Count	Percent	Count	Percent	Count	Percent
Balanced	17	1 %	1	4 %	0	0 %
Buyout	305	21 %	10	38 %	3	16 %
Co-investment	4	0 %	1	4 %	0	0 %
Distressed Debt	36	2 %	0	0 %	0	0 %
Fund of Funds	92	6 %	0	0 %	0	0 %
Growth	36	2 %	1	4 %	7	37 %
Infrastructure	11	1 %	0	0 %	4	21 %
Mezzanine	64	4 %	0	0 %	1	5 %
Natural Resources	34	2 %	1	4 %	0	0 %
Real Estate	351	24 %	4	15 %	3	16 %
Secondaries	4	0 %	0	0 %	0	0 %
Special Situations	12	1 %	0	0 %	0	0 %
Timber	7	0 %	0	0 %	0	0 %
Turnaround	4	0 %	1	4 %	0	0 %
Venture	481	33 %	7	27 %	1	5 %
Size						
Small	320	22 %	15	58 %	6	32 %
Medium	731	50 %	9	35 %	9	47 %
Large	403	28 %	2	8 %	4	21 %
Decade						
90s	570	39 %	10	38 %	2	11 %
00s	888	61 %	16	62 %	17	89 %
Sequence Number						
0	70	5 %	3	12 %	0	0 %
1	301	21 %	8	31 %	9	47 %
2	317	22 %	5	19 %	6	32 %
3	221	15 %	4	15 %	2	11 %
4	549	38 %	6	23 %	2	11 %
Cycle						
Boom	1237	85 %	23	88 %	14	74 %
Bust	221	15 %	3	12 %	5	26 %

Table 24 - Continent Comparison by Type, Size, Decade, Sequence Number and Cycle

Comparrisons of All Fund Types							
		IRR			TVPI		
Fund Type	Obs	Median	P-Value	Test	Median	P-value	Test
Buyout & Venture	786	8.6			1.46		
Balanced	17	13.1	0.2543	MWW_2	1.85	1.3799	MWW_2
Co-investment	4	6.6	0.9553	MWW_2	1.355	0.9641	MWW_2
Distressed Debt	36	16.6	0.0040	MWW_2	1.56	1.8525	MWW_2
Fund of Funds	92	8.8	0.3402	MWW_2	1.45	1.4240	MWW_2
Growth	36	10.4	0.3608	MWW_2	1.665	1.8303	MWW_2
Infrastructure	11	14.4	0.4361	MWW_2	1.62	1.0940	MWW_2
Mezzanine	64	9.55	0.8152	MWW_2	1.38	0.7978	MWW_2
Natural Resources	34	20.55	0.0000	MWW_2	1.835	1.2484	MWW_2
Real Estate	351	9.3	0.5801	MWW_2	1.4	0.0937	MWW_2
Secondaries	4	10.65	0.6580	MWW_2	1.5	0.8917	MWW_2
Special Situations	12	10.15	0.6204	MWW_2	1.565	1.4628	MWW_2
Timber	7	1.9	0.1047	MWW_2	1.42	0.8985	MWW_2
Turnaround	4	6.8	0.9562	MWW_2	1.305	0.6328	MWW_2

Table 25 - Comparison of Fund Types Outside Our Sample

Appendix H – Quartile Tables

Top Quartile in Consecutive Periods: All Buyout and Venture									
IRR						Last Funds Quartile Rank			
1st Quartile	Count	Median	Mean	Std Error	Conf Int	1st	2nd	3rd	4th
1 period	147	10.9	26.6626	4.7906	9.3895		0.2618	0.0108	0.0000
2 perods	36	20.05	42.3028	14.5717	28.5599	0.1240	0.0171	0.0041	0.0000
3 periods	11	23.3	31.5000	10.2913	20.1705	0.0901	0.0288	0.0151	0.0008
TVPI						Last Funds Quartile Rank			
1st Quartile	Count	Median	Mean	Std Error	Conf Int	1st	2nd	3rd	4th
1 period	147	1.67	2.5281	0.2679	0.5250		0.1789	0.0035	0.0000
2 perods	36	2.19	3.1994	0.6677	1.3086	0.1435	0.0247	0.0021	0.0000
3 periods	11	2.28	2.2755	0.3840	0.7525	0.3115	0.1118	0.0329	0.0032

Table 26 - Comparison of Consecutive Top Performing GPs

Comparison of Last Quartile Rank for Different Characteristics							
US Buyout and Venture Funds - Last Funds' Quartile Performance							
IRR	2nd	3rd	4th	TVPI	2nd	3rd	4th
1st	0.3630	0.0169	0.0000	1st	0.3437	0.0132	0.0000
2nd		0.0736	0.0000	2nd		0.1079	0.0000
3rd			0.0013	3rd			0.0010
All Buyout and Venture Funds - Last Funds' Quartile Performance							
IRR	2nd	3rd	4th	TVPI	2nd	3rd	4th
1st	0.2618	0.0108	0.0000	1st	0.1789	0.0035	0.0000
2nd		0.1013	0.0000	2nd		0.1079	0.0000
3rd			0.0024	3rd			0.0026
US All Funds - Last Funds' Quartile Performance							
IRR	2nd	3rd	4th	TVPI	2nd	3rd	4th
1st	0.9672	0.0268	0.0000	1st	0.7250	0.0539	0.0000
2nd		0.0065	0.0000	2nd		0.0910	0.0000
3rd			0.0001	3rd			0.0000
All Funds - Last Funds' Quartile Performance							
IRR	2nd	3rd	4th	TVPI	2nd	3rd	4th
1st	0.7692	0.0131	0.0000	1st	0.5220	0.0166	0.0000
2nd		0.0124	0.0000	2nd		0.0662	0.0000
3rd			0.0001	3rd			0.0001

Table 27 - Comparison of Last Funds' Quartile Performance by Included Datapoints

Next Fund Top Quartile Chance		
Top Quartile in	Obs	Next Top
1 period	886	34.88 %
2 periods	193	40.93 %
3 periods	49	59.18 %

Table 28 - Percentage of Next Funds in Top Quartile based on Top Quartile Persistance

Percentage of Funds With An IRR Above				
Top quartile in	Obs	Benchmark	20.1	0
1 period	356	59.27 %	33.43 %	90.37 %
2 periods	78	62.82 %	35.90 %	90.45 %
3 periods	25	80.00 %	56.00 %	93.94 %

Table 29 - Percentage of Next Fund Beating Benchmark, Top Quartile and Gaining Positive Results by Top Quartile Persistance

Appendix I – MWW-test A Numerical Example

MWW-test - Numerical Example						
Data samples		Ranks		Statistics		
Male	Female	Male	Female		Male	Female
179	165	12	2.5	n	10	10
186	171	17.5	7	Median	182.5	170.5
165	180	2.5	13	Rank sum	140.5	69.5
193	163	20	1	U	14.5	85.5
189	172	19	8			
182	169	15	5	U - min	14.5	
177	170	10	6	Mean U	50	
183	168	16	4	Variance	175	
186	173	17.5	9	Std dev	13.23	
178	181	11	14	z-score	-2.6835	

Table 30 - MWW test - A numerical example

In the two leftmost columns in Table 30, we have two samples showing the heights (cm) of 10 men and 10 women. In the “Ranks” columns, we have given each observation, independent of samples, a rank in an ascending order. As we see from the table, the lowest (163 cm) out of the 20 observations has gotten the rank 1, and the tallest (193 cm) has gotten the rank 20. In the event of equal observations, also called a tie, the average rank for the two observations will be assigned both of them. By summing the ranks, the rank sum (R) is found for each sample. Now that the R is found, utilising Formula 9 and 10 will give the test score.