Erasmus School of Economics, Erasmus University Rotterdam

Does fund size erode mutual fund performance?

An estimation of the relationship between fund size and fund performance

In this paper I try to find the true relation between the size of a mutual fund and its performance. I try to find this relationship by answering the main research question: does fund size erode mutual fund performance? This main question is split into three different hypotheses, which all look at different sides of this relationship. The main findings of this paper are that there indeed is a negative relation between fund size and its performance, that there is not enough evidence to support a positive relationship between fund size and its performance within fund families and that there does indeed seem to be an optimal size for a mutual fund and that this optimum is somewhere in the range of $18,6 and $58,8 million net assets under management. This paper is of interest to both individual investors because it might influence their investment strategy and to fund managers because it might influence their fund management when it comes to the structuring and daily operations of a mutual fund.

Jeroen van Geel, studentnumber 367547
18-7-2016

Table of Contents

I. Introduction

The mutual fund industry has been booming business over the past two decades. Since the early 1980's billions, and nowadays, even trillions of dollars have been invested in all types of mutual funds. Looking at the figures from 2014, it can be seen that the U.S. mutual fund industry remains the largest in the world with almost 16 trillion dollars in assets. By far the most popular mutual fund remains the Equity mutual funds, which held over half (52%) of the total mutual fund assets. Mutual funds are not only very popular with institutional investors but also with everyday households. In fact, the everyday households are the largest holders of ownership in mutual funds (Investment Company Institute, 2015).

The largest part of this very fast growing industry consists of funds that invest in stocks and are actively managed. Investors therefore need lots and lots of information to identify managers with stock-picking ability. This is part of the reason of the explosion of information provided in newsletters, magazines and rating services like Morningstar. With funds growing as big as they have in recent years they control a very sizeable amount of the corporate equity market and play a pivotal role in the determination of stock prices.

With mutual funds growing bigger and bigger, one might ask if bigger is better? Are there economies of scale in the mutual fund industry? On the one hand the argument of lowering transaction costs and being able to invest in smaller less liquid firms can lead to a decrease in costs and an increase in investment opportunities, both of which can be marked as positives. On the other hand, a larger institution can lead to agency problems and organizational diseconomies. Bigger funds also run into the problem described in the paragraph above, they influence the determination of prices. When this price impact becomes too big it might eat into the profits that can be made from investments that otherwise for smaller firms would be very profitable.

In this paper, I try to get a better understanding of how the economies of scale work in equity funds. With the seemingly endless growth of investments in mutual funds of all sorts and types, it is possibly more crucial than ever to establish the relationship between size and performance. There might be a good reason to not invest in the bigger and bigger growing mutual funds if there is a negative relationship. This paper will try to find an answer to the following question:

Does fund size erode mutual fund performance?

This paper will follow the paper by Chen et. Al. (2004) closely. The main difference will be the extension of the data used, in their paper, a time period up until 1999 is considered while this paper will look at the period after that, 1999 till 2015 to be exact. This paper will also split the main research question into 3 hypotheses. The first one will look at the relation between size and performance on different levels. It will use month and firm fixed effect regressions to examine the size effect in both a time-series as in a cross-sectional setting. The first hypothesis will also be extended by running the regression for each decile to see what the difference for each decile are. The second hypothesis will look at the effects within family funds. The regression will use interaction dummies to only examine the relationship within family funds for each month. The third and last hypothesis will try to find the optimal size of a fund. With the estimated coefficients, descriptive statistics and economical reasoning an optimal size will be determined for both the return side as well as the cost side of mutual funds.

The primary findings of this paper are that the size effect is observable with mutual funds in the chosen time period. The fixed effects regressions also show that this effect is also observable in both a time-series as in a cross-sectional setting. The size effect within fund families does seem to exist as well, meaning that also within fund families being a bigger fund leads to lower performance. The other finding that came from the second hypothesis is the fact that a bigger fund family does not lead to higher performance. In fact, the opposite is true, a bigger fund family leads to lower performance. the optimal size of a fund seems to be somewhere in the 5^{th} or 6^{th} size decile. This is true for both the return as well as the cost side of mutual funds.

Next, I will continue with a review of existing literature about the economies and diseconomies of scale in mutual funds. After that, I will break down the research question into three individual hypotheses and explain how they contribute to find a good answer on the research question. The fourth part of this paper will explain what data is used and where it was obtained from. It will also give some generic statistics about the dataset. The fifth part will explain the used methodology in more detail. With the data collected and the methodology determined, the sixth part of this paper will report and interpret the results. The last part of this paper will consist of a short summary of the research, some concluding remarks and some points of interest for future research.

II. Literature review

Even though the growth of the mutual fund industry has been incredible and the relation between size and performance seems an important one, not a lot of research has been done in the field of economies of scale in the mutual fund industry. A common view by practitioners is that a larger fund has a number of advantages. The most significant being the availability of more resources for research and lower expense ratios. However, others believe that a large asset base can erode fund profits because of trading costs associated with liquidity or price impact (André Perold and Robert S. Salomon (1991); Roger Lowenstein (1997)). A small fund can easily put all or a big amount of its assets in its best ideas, a lack of liquidity forces a large fund to invest in the not-so-good investment ideas or take larger than optimal positions per stock. This leads to erosion of the profits that are made by large funds. Grinblatt and Sheridan Titman (1995) find mixed evidence that fund returns decline with fund size for a small sample of funds in the period 1974 to 1984. A lot of the previous research has been focused on the persistence of mutual fund performance and the difference between actively managed and passive funds. The next paragraphs will give a short summary of previous research.

Such a large player in the financial market can be a good way to learn about investment strategies and how asset prices are determined and influenced by the players on the market. Since mutual funds have been growing at a very high pace, their performance and the difference between funds has been researched quite a bit. Research done by Cuthbertson, Nitzsche and O'Sullivan (2010) shows that there are little to no equity mutual funds in the US and the UK which actually record a positive-alpha performance (0-5% of top performing funds). Most of the time the positive returns made by a lot of funds are lost in load fees, expenses and turnovers. Their research also shows that there is little evidence for successful market timing and that effectively 75% of the active funds they study are zero-alpha funds (Cuthbertson, Nitzsche, & O'Sullivan, 2010). Even though it seems that investing in active funds seems to serve no purpose, Grinblatt and Titman (1992) show that there is a difference in performance between funds over time and that this difference is persistent and can be attributed to the ability of fund managers to get abnormal returns (Grinblatt & Titman, 1992).

There seems to be no clear consensus as to whether it is better to invest in index funds or in actively managed funds because in a paper by Fortin and Stuart (1999) it becomes clear that in 25 out of the 30 tested cases, index funds outperform actively managed funds. The only difference can be seen when the researchers look at funds that invest in small cap. There, the actively managed funds outperform the index 4 out of 5 times. The researchers think this is due to the fact that there are more equity market inefficiencies in the small cap market which fund managers can exploit to beat the index (Fortin & Michelson, 1999). So if actively managed funds tend to underperform simple index funds, why is there still a large growth in investments in actively managed funds? Grinblatt (1996) tries to find an answer to this question and comes to a surprising conclusion. He says that future fund performance is predictable by looking at past performance. This can be the case because the price at which funds are bought and sold is equal to net asset value and does not change to reflect superior management. This means that there are no extra costs to having superior management (Grinblatt M. , 1996).

The most extensive research comes from Chen, Hong, Huang and Kubik (2004), who not only look at the relationship between size and fund performance but also at how liquidity and the organizational structure of the funds, or the fund family, influence the performance of mutual funds. They find a negatively relationship between size and fund performance both before and after fees and expenses with individual firms. When looking at funds that are part of a fund family size does not need to have a negative influence on fund performance. The researchers think this is because fund managers in a fund family have their own pile of money to invest and do not have to compete with other managers in order to receive the same amount of money next year. They also look at organizational diseconomies and find that those also negatively influence fund performance (Chen, Hong, Hiang, & Kubik, 2004). Research done before this have shown that there seems to be a minimum size fund in order to achieve sufficient returns to justify their costs of acquiring and trading on information. But, that there are diminishing marginal returns to information acquisition and trading and the marginal returns become negative when the mutual fund exceeds its optimal fund size (Indro, Jiang, Hu, & Lee, 1999). On the other hand, Reuter and Zitzewitz (2010) show that any downward bias in standard estimates of performance persistence due to diseconomies of scale is likely to be small (Reuter & Zitzewitz, 2010).

III. Hypotheses

The proposed research question is:

Does fund size erode mutual fund performance?

This is a pretty broad question and takes a lot of different aspects of mutual funds in consideration at once. For this reason, I split this research question into smaller hypotheses. These will look at different aspects of mutual funds and how the mechanics of economies of scale work in this industry. The results found in these hypotheses will contribute towards the final answer of the research question. The next part will explain what the hypotheses are and how they contribute towards finding an answer to the research question.

The first hypothesis is as follow:

H_1 : There is a negative relationship between the size of a mutual fund and its performance

This first hypothesis is a test of the most basic relationship between size and performance of mutual funds. It is there to have a starting point for further investigation. I expect this hypothesis to be true in accordance with the result found in the paper of Chen et. Al. (2004). If this hypothesis is accepted, it becomes clear that simply putting your money in big and established funds is not necessarily better than picking smaller funds that have good track records. It also provides an opportunity for more competition between funds. Once a certain size has been reached any excess growth would result in negative returns. It would therefore be more logical to start a new fund, which would create more competition and maybe reduce load fees and expense ratios.

One thing to take into account with this first hypothesis is that it might be a little too basic to define the true relationship between size and performance. As will be explained in the next part of this paper, it will control for other variables influencing the performance of a fund it but it cannot take all effects into account. A second consideration would be the fact that this hypothesis only looks at the profit side of a mutual fund and not so much at the costs a fund has to deal with. The last consideration is the fact that the dataset is panel data and to examine both the time-series and cross-sectional settings for this data set, I will include fixed effect regressions to account for these different sides of the size effect. I will run the regressions per size decile to get an even clearer view of the size effect.

To extend the first hypothesis, I will look at family funds for the second hypothesis. The point of interest for this hypothesis is to see if there are also decreasing returns to scale within fund families. The hypothesis is specified as follows:

H_2 : There is a positive relationship between the size of a fund within a family and its return

This hypothesis tests the second result from Chen et. Al. (2004), which states that being part of a larger family fund is better, at least for most of the time, than being a large stand-alone fund. I expect this hypothesis to be true as well. Implications would be that the optimal way to structure a mutual fund would be to have one big overseeing fund which has numerous smaller funds beneath itself that individually invest. The individually invest part of the last sentence is very important because, like it was stated by Chen et. Al. (2004), one possible explanation for this phenomenon could be the fact that fund managers do not have to compete for funds on a monthly basis when funds are part of a larger family. The most observed structure in family funds have a set allocation per fund, of course this can change over time but it will not change as much as with stand-alone firms. Another part of this hypothesis test if there are diseconomies of scales within fund families. It could be the case that within a fund family a large set of smaller funds are better than having a smaller number of bigger funds per family. I do not really know what to expect for this second part of the hypothesis. On the one hand I expect a bigger fund within a family to be able to lower the costs of management because it can be spread those costs within the family, the bigger fund can therefore report higher returns than the smaller firms. But it can also be the case that a bigger fund within a family comes across the same diseconomies of scale as a stand-alone fund.

The third and last topic of this paper tries to find an optimum size for a mutual fund. Like it was stated by Indro et. Al. (1999), there seems to be an optimum size at which a mutual fund can operate. There seems to be a minimum size because a mutual fund has to be big enough to justify the cost of acquiring and trading on information and gain a sufficient return on that information. But on the upper bound the marginal returns from this information are diminishing and so is trading on this information, so there seems to be a maximum size as well. To find this optimum fund size, I will look at the expense ratio and other costs, like management fees and operating expenses, of these mutual funds per size decile and look for any patterns. When the optimal size decile for costs is determined it is of course also important to look at the returns side. I expect that for the costs the rule of "Bigger is better" applies because a bigger fund can spread its costs more across more investments and customers and can therefore lower the costs as a whole. When it comes to returns the relationship is not so clear I think. Bigger could be better as a bigger fund could invest in much more instruments and markets and can maybe get discounts when buying new stocks. But it also means that a very big fund can run into liquidity issues and influence the market price too much.

IV. Data

The data to be used to try and find an answer to this research question will be obtained through a similar way as done by Chen et. Al. (2004). The Center for research in Security Prices (CRSP) Mutual Fund Database will provide the information about the mutual funds. This includes the size (measured by total assets under management), monthly returns, expense ratio, the age of the fund, the monthly turnover, the total load charged and if the fund is part of a fund family the size of the family is included as well. With the fund characteristics we can control for spurious effects in the relationship between size and performance. With the inclusion of the size of the fund family we can look if the result found in Chen et. Al. (2004) are also found in the period after their research. Only diversified US equity mutual funds are used in this paper. For a fund to enter the sample, it must report information on assets under management and monthly returns. In order to create a benchmark performance for each fund, every fund has to have at least one year of reported returns. This excludes very young funds from the sample.

The data will be split into size deciles based on their previous month TNA to get a more detailed view of the true relationship between size and fund performance. Another advantage of splitting the funds into size deciles is the fact that effects in deciles can be explored further if need be. This is also helpful with the third and last part of the paper which focusses on determining an optimal size for a mutual fund.

Fund returns are measured both in gross returns and in net returns and adjusted for the market return. The market return used here is the Value-Weighted return including dividend from the CRSP Stock Market Index. Net returns are the returns in excess of fund expenses, like management fees, and load fees. A fund can report abnormal returns on a year to year basis but if the excess return after expenses is not higher than the index, an investor would have no reason to invest through a mutual fund as oppose to just buying the index. To calculate the Net fund, return for a specific month for a specific fund, I take the reported return from the CRSP database and subtract the market return. To calculate the gross fund return, I divide the quarterly expense ratio as reported by the mutual funds in the CRSP database by three and add those to the monthly net fund returns.

I restrict the amount of mutual funds to those firms that are traded on the American NASDAQ stock exchange. The second restriction is based on the Lipper Asset Classification Code. The classification code is based on how a mutual fund describes the way it intends to trade in its prospectus, it also examines the holdings of the funds and measure them to a benchmark portfolio and then classifies the funds per asset class (Center for Research in Security Prices). For this paper I am only interested in Equity mutual funds, therefore I select only those funds that are classified as “EQ”. The next step is to select only those funds that report returns on a monthly basis. In order to form a benchmark for the performance of each fund it has to have reported at least one year of returns, I therefore eliminate all firms that have less than 12 reported months of returns. With these restrictions a mutual fund can enter the database multiple times per month as it may invest in different share classes. I clean the data by eliminating those redundant observations. What I end up with is a dataset consisting of 2133 different funds. This is a smaller number than the 3439 used by Chen et. Al. (2004).

Descriptive statistics for all variables can be found in tables 1 through 4, found below. In each month there are on average 1030 funds in my sample. These funds have an average total net assets (TNA) of $322,6 million, with a standard deviation of $1663 million. As can be seen from the standard deviation, there seems to be a large spread in the size of the fund's net assets in the sample. If we look at table 4 we can see where this arises from as the smallest 10% of funds only have an average of $0,51 million TNA and the biggest 10% of funds have $2546,7 million TNA on average. This difference is very significant. Besides looking at the TNA of each fund in the sample, the family fund size could also be important in the relation between size and performance. I therefore include the logarithm of the total net assets of the family each fund is part of.

Besides the size of the funds and their fund family size the CRSP database has a wide variety of fund statistics that will be used as control variables in the estimation of the relationship between size and performance. These include: the age of a fund, the turnover ratio, the expense ratio and the management fees made by the funds. Averages, standard deviations of these averages and the minimum and maximum values for all these variables can be found in table 1 for the total sample and table 4 for a breakdown per size decile. Table 2 reports the cross-sectional averages of the correlations between the variables and Table 3 reports the average gross and net fund return per size decile.

Table 1 Descriptive statistics of total sample			
	Mean	Minimum	Maximum
Number of Funds	1030,20	-	-
TNA	322,55	0,00	73313,50
($ million)	[1662,99]		
LOGTNA	3,40	-2,30	11,20
($ million)	[2,37]		
LOGFAMSIZE	4,07	-2,30	14,12
($ million)	[2,36]		
AGE	14,67	0,83	87,75
(years)	[9,77]		
TURNOVER	1,01	0,00	843,87
(% per year)	[5,90]		
EXPRATIO	0,02	0,00	1,46
(% per year)	[0,02]		
MGMTFEE	0,85	0,00	44,30
(% per year)	[0,67]		

Table 1 Descriptive statistics for the used variables in the data set

Table 2 Cross-sectional averages of correlation between variables								
	TNA	LOGTNA	LOGFAMSIZE	AGE	TURNOVER	EXPRATIO	MGMTFEE	ACTUAL12b-1
TNA	1							
LOGTNA	0,44	1,00						
LOGFAMSIZE	0,32	0,63	1,00					
AGE	0,31	0,32	0,19	1,00				
TURNOVER	-0,03	-0,09	-0,13	-0,12	1,00			
EXPRATIO	-0,12	-0,44	-0,41	-0,06	0,12	1,00		
MGMTFEE	0,00	0,11	0,05	-0,06	0,14	0,28	1,00	

Table 1 Cross-sectional averages of correlation between all used control variables

Table 3 Average Net and Gross return per decile and for the total sample											
	Decile 1	Decile 2	decile 3	Decile4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Total
FUNDRET	0,67%	0,58%	0,52%	0,52%	0,43%	0,48%	0,44%	0,40%	0,35%	0,28%	0,453%
(Gross)	4,43%	3,63%	3,35%	3,41%	3,35%	3,37%	3,33%	3,38%	3,10%	3,00%	3,45%
FUNDRET	-0,194%	-0,089%	-0,119%	-0,042%	-0,096%	-0,019%	-0,032%	-0,038%	-0,068%	-0,078%	-0,085%
(Net)	4,00%	3,55%	3,48%	3,48%	3,38%	3,38%	3,30%	3,34%	3,10%	2,99%	3,44%

Table 2 Average net and gross fund returns per size decile and for the total sample. The net fund returns are the returns reported by the mutual funds from the CRSP database and are the returns after all expenses. The gross fund returns are the net fund returns plus the total expenses reported by the mutual funds

Table 4 Descriptive statistics per size decile

	Decile 1			Decile 2			Decile 3			Decile 4			Decile 5			Decile 6			Decile 7			Decile 8			Decile 9			Decile 10		
	mean	min	max	mean	min	max	mean	min	max	mean	min	max	mean	min	max	mean	min	max	mean	min	max	mean	min	max	mean	min	max	mean	min	max
Number of Funds	101,15	-	-	99,58	-	-	99,44	-	-	100,76	-	-	99,42	-	-	100,05	-	-	99,94	-	-	99,90	-	-	99,98	-	-	100,01	-	-
LAGGEDTNA	0,51	0,00	1,40	2,92	1,50	4,60	7,17	4,70	10,10	14,04	10,20	18,50	25,00	18,60	33,20	45,23	33,30	58,80	77,88	58,90	103,10	150,61	103,20	214,10	354,28	214,20	598,70	2546,73	598,80	73313,50
($ million)	0,42			0,90			1,60			2,42			4,20			7,42			12,69			31,83			107,25			4688,97		
LOGTNA	-1,10	-2,30	0,34	1,02	0,41	1,53	1,95	1,55	2,31	2,63	2,32	2,92	3,20	2,92	3,50	3,80	3,51	4,07	4,34	4,08	4,64	4,99	4,64	5,37	5,83	5,37	6,39	7,36	6,39	11,20
($ million)	0,98			0,32			0,23			0,17			0,17			0,17			0,16			0,21			0,29			0,82		
LOGFAMSIZE	1,37	-2,30	12,19	2,24	-2,30	13,28	2,81	-2,30	12,09	3,29	-2,30	11,26	3,94	-0,51	11,85	4,46	-1,20	12,47	5,01	-2,30	12,51	5,64	-2,30	14,12	6,43	3,50	13,83	7,97	1,44	13,45
($ million)	2,75			1,96			1,53			1,36			1,37			1,20			1,17			1,11			1,07			1,09		
AGE	8,79	0,83	51,08	11,03	0,83	80,25	11,97	0,83	80,25	13,24	0,92	80,25	13,62	0,92	80,25	14,97	0,92	80,25	16,83	1,25	80,25	17,19	1,25	87,75	18,34	1,33	87,75	21,76	1,00	72,67
(years)	6,15			6,94			8,31			9,57			8,92			8,90			9,97			9,80			10,45			10,71		
TURNOVER	1,98	0,00	843,87	1,27	0,00	85,76	1,33	0,00	231,17	1,09	0,00	107,11	1,00	0,00	34,52	0,94	0,00	60,12	0,82	0,00	60,12	0,82	0,00	24,31	0,71	0,00	9,25	0,53	0,00	7,06
(% per year)	19,19			3,69			6,76			3,57			1,40			1,66			1,40			1,12			0,85			0,64		
EXPRATIO	0,03	0,00	1,46	0,02	0,00	0,15	0,02	0,00	0,07	0,02	0,00	0,06	0,02	0,00	0,06	0,01	0,00	0,06	0,01	0,00	0,06	0,01	0,00	0,04	0,01	0,00	0,03	0,01	0,00	0,03
(% per year)	0,08			0,01			0,01			0,01			0,01			0,01			0,00			0,00			0,00			0,00		
MGMTFEE	0,92	0,00	44,30	0,86	0,00	4,73	0,85	0,00	4,87	0,79	0,00	36,11	0,83	0,00	4,82	0,86	0,00	3,08	0,87	0,00	4,82	0,86	0,00	2,56	0,86	0,00	2,78	0,78	0,00	2,16
(% per year)	1,99			0,49			0,52			0,81			0,42			0,35			0,33			0,31			0,32			0,31		

Table 4 descriptive statistics for all used control variables split into size deciles for a more detailed analysis. The number of funds in a decile is the average amount of funds in a decile for the entire sample and does therefore not have to be a whole number.

Looking at these statistics a couple of patterns arise. For example, the bigger the fund the older it is on average. This makes sense as funds that perform well get more funds and grow bigger which ensures their sustainability and therefore older firms tend to be bigger than smaller funds. Another pattern is the fact that the bigger the fund the lower its turnover ratio is. It seems to be the case that smaller funds move their money more often than larger funds. This may be related to the investment strategy difference between small and big funds. Smaller funds tend to invest in smaller companies and try to find the big winners of the future, whilst bigger funds tend to diversify more and put their money in safe bets that pay regular income streams. Smaller firms are more volatile than bigger firms and therefore small funds have to move their money more often than bigger funds. The last interesting pattern is the pattern in the management fees paid by the funds. For the lower size deciles (1 through 4) this percentage seems to be decreasing but for deciles 5 through 7 this percentage increases a bit again. After decile 7 the management fee drops considerably. It might be the case that the actual level of the management fees stays the same throughout the entire sample but that it drops as a percentage of the size or it might be the case that larger funds actually have lower management fees that smaller firms.

In table 2 the cross-sectional averages of the correlations between the different variables can be found. Interesting things to note from this table are the correlations between TNA, LOGTNA and LOGFAMSIZE. The correlations do not seem to be very high, this would suggest that there is not a large effect on the size of an individual fund and the size of the family it is part of. Another set of interesting correlations are the correlations between age and the “costs” variables. These correlations are all negative which would suggest that the costs a fund has decrease as a fund gets older. The last set of interesting correlations are the correlations between LOGFAMSIZE and the “cost” variables. These are also negative, except for the management fee variable, this would suggest that being part of a bigger fund family leads to less costs.

In table 3 the averages of the Gross and Net fund returns can be seen for both the entire sample and broken down into the different size decile portfolios. To get the Net fund returns, I have subtracted the Value-Weighted return including dividends from the returns as reported by each fund. To get the Gross return, I divided the quarterly reported expense ratio by three to get the monthly expense ratio and added that to the net fund return as calculated above. For the Gross returns it seems to be the case that the highest decile funds do considerably worse than the lowest decile funds. This is what I expect to see when there are decreasing returns to scale for mutual funds. The total average Gross return is 0,45%. Looking at the Net return the pattern seems to be the other way round, the largest funds do better (or less bad because all the returns are negative) than the smallest funds. the total average Net return is -0,09%. Interesting thing to note is that none of the ten decile portfolios have a positive Net return which would mean that they do not perform better than the market portfolio.

V. Methodology

Using cross-sectional variation between firms leads to some concerns. The main concern is the fact that there could be heterogeneity between firm styles across firm sizes. A small fund might be more likely to pursue small stock, value stock and price momentum strategies than a large fund. Small stocks tend to be more volatile than large stocks and therefore end up with higher returns. If this heterogeneity needs to be accounted for is arguable but it would be nice if I could end up with results that indicate that fund size influences fund performance even after accounting for variations in fund styles. To account for this heterogeneity in fund styles, I adjust the fund performance by different benchmarks. Besides the simple market-adjusted return benchmark I include the Capital Asset Pricing Model (CAPM) of William F. Sharpe (1964), the three-factor model of F. Fama and Kenneth R. French (1993) and the augmented three-factor model which includes the momentum factor of Narasimhan Jegadeesh and Titman (1993). These models have been proven to have additional explanatory power for the observed cross-sectional variation in fund performance for example by Carhart (1997).

Table 5 reports descriptive statistics of these factors that make up my portfolio benchmarks. Included in these factors are the CRSP value weighted stock index return net of the one-month Treasury rate (VWRF), the returns to the Fama and French (1993) Small stocks minus Big stocks (SMB) and High book-to-market stocks minus low book-to-market stocks (HML) portfolios, the last factor reported is the Jegadeesh and Titman (1993) returns-to-price momentum portfolio (MOM12). This last factor consists of a portfolio of long positions in the past-12-months winners and short positions in the past-12-months losers and hold for one month.

Table 5 Summary statistics Factors

			Cross-correlations			
Factor	Mean return	SD of return	VWRF	SMB	HML	MOM12
VWRF	0,38%	4,55%	1,00			
SMB	0,37%	3,56%	0,28	1,00		
HML	0,19%	3,38%	-0,16	-0,35	1,00	
MOM12	0,38%	5,71%	-0,33	0,12	-0,18	1,00

Table 5 descriptive statistics for the factors used in calculating the adjusted returns and their cross-correlations

The goal of this paper is to find the relationship between fund size and fund performance, I therefore sort the funds at the beginning of each month on their previous month TNA into ten decile portfolios. After that I follow these portfolios for one month and use the entire time series of their monthly net returns to calculate the loadings to the various factors (VWRF, SMB, HML and MOM12) for each portfolio. I then assign each of these loadings to the funds based on their size. This means for example that when a portfolio stays in the same-size decile for the entire sample that its loadings to the above factors stays the same and that when it moves from one size decile to the next size decile it gets the loadings of those decile with which the next month's return for that fund will be adjusted.

Table 6 reports the loadings of the ten fund-size (TNA) sorted mutual fund portfolios using the CAPM:

$$(1)\ R_{i,t} = \alpha_i + \beta_i VWRF_t + \varepsilon_{i,t}$$

Where $R_{i,t}$ is the (net fund) return on one of the ten fund-size mutual fund portfolios in month t in excess of the one-month T-bill return, α_i is the excess return of that portfolio, β_i is the loading on the market portfolio and $\varepsilon_{i,t}$ stands for the generic error term that is uncorrelated with all the other independent variables. Other papers have found that, on average, a mutual fund has a beta of around 0.91 this reflects the fact that mutual funds hold some cash or bonds in their portfolios. The beta's found here are consistent with this finding and are consistent with Chen et. Al. (2004). Looking at the estimated alphas for this model something interesting becomes apparent. The alpha for the smallest fund decile is lower than the alpha for the highest fund decile and that for most of the deciles there seems to be an upward trend in the alpha's. This is opposite of the expectation of decreasing return to scale, since if there are decreasing return to scales these alphas should be declining the larger the decile. One possible solution for this might be the fact that the time period of interest for this paper includes the financial crisis. In this financial the crises the biggest funds might have survived based off of their reserves being higher and their ability to invest in top tier safe stocks. The smaller funds have less reserves and tend to invest more in smaller companies and therefore their investments are riskier than those of big funds. it would only make sense for the smallest funds to go bankrupt more easily than bigger funds and this might be a possible reason for the fact that these alphas' do not show decreasing return to scales.

Table 6 Loadings calculated using the CAPM

	CAPM	
Portfolio	Alpha	VWRF
1 (small)	0,02%	0,91
2	0,11%	0,93
3	0,07%	0,91
4	0,14%	0,92
5	0,09%	0,93
6	0,19%	0,92
7	0,16%	0,94
8	0,15%	0,97
9	0,14%	0,95
10 (large)	0,17%	0,90

Table 6 estimated loadings calculated using the CAPM model. The estimated VWRF values are commonly referred to as the Beta for a company. This beta measures the sensitivity of a company to the market portfolio. The Alpha in this model can be interpreted as a measure of excess return

Table 7 reports the loadings for the two additional performance models the three-factor model of Fama and French (1993) and this augmented three-factor model with the momentum factor:

$$(2)\ R_{i,t} = \alpha_i + \beta_{i,1}VWRF_t + \beta_{i,2}SMB_t + \beta_{i,3}HML_t + \varepsilon_{i,t}$$
$$(3)\ R_{i,t} = \alpha_i + \beta_{i,1}VWRF_t + \beta_{i,2}SMB_t + \beta_{i,3}HML_t + \beta_{i,4}MOM12_t + \varepsilon_{i,t}$$

Where $R_{i,t}$ and α_i represent the same underlying variables as in formulae (1), the β_i's are loadings on the various benchmark portfolio's and $\varepsilon_{i,t}$ stands again for a generic error term that is uncorrelated with all the other independent variables. When looking at the estimated alpha's again, we can see that there still is an upward trend overall but there is not a clear increase or decrease across all deciles. From the table we can see that there are no clear patterns in the loadings for SMB and HML factors in both the three-factor and the augmented three-factor model. This is surprising as Chen et. Al. (2004) find that smaller funds tend to have a higher loading for both of these factors than bigger funds. A possible explanation for this is the fact that their sample was split into quantiles and not in deciles, with deciles we can see the true loadings in more detail. The loadings of the momentum factor on the other hand show an increasing pattern with fund size. Another interesting thing to note from the loadings to the market factor (VWRF) is the fact that they all have declined when compared to the CAPM model. This is consistent with Carhart (1997) who found that the three-factor model has additional explanatory power in mutual fund performance.

Table 7 Loadings calculated using the 3-Factor model and the 4-Factor model

	3-Factor model				4-Factor model				
Portfolio	Alpha	VWRF	SMB	HML	Alpha	VWRF	SMB	HML	MOM12
1 (small)	-0,02%	0,88	0,14	0,05	-0,02%	0,87	0,15	0,04	-0,02
2	0,06%	0,90	0,16	0,08	0,06%	0,89	0,16	0,07	-0,02
3	0,04%	0,90	0,10	0,07	0,04%	0,88	0,11	0,06	-0,04
4	0,10%	0,90	0,13	0,05	0,10%	0,89	0,13	0,04	-0,01
5	0,03%	0,90	0,18	0,03	0,03%	0,90	0,18	0,03	0,01
6	0,12%	0,90	0,16	0,09	0,12%	0,90	0,16	0,09	-0,01
7	0,10%	0,92	0,15	0,07	0,10%	0,92	0,15	0,07	0,00
8	0,08%	0,94	0,18	0,03	0,08%	0,95	0,18	0,04	0,02
9	0,07%	0,91	0,19	0,06	0,07%	0,91	0,19	0,06	0,01
10 (large)	0,12%	0,89	0,10	0,07	0,12%	0,89	0,10	0,07	0,01

Table 7 estimated loadings calculated using both the 3-factor model and the 4-factor model. These loadings are used to calculate the predicted returns for each decile in each month

Now that the loadings for each benchmark are calculated I can continue with adjusting the return of each fund and find the adjusted return that will be used to specify the relation between fund size and fund performance. Because I want this relation and only this relation, I include other fund characteristics which might have some explanatory power in fund performance. With this I can account for any spurious effects that might arise. The specific regression I run is the following:

$$(4)\ Fundreturn_{i,t} = \alpha_i + \beta_i LOGTNA_{i,t-1} + \gamma X_{i,t-1} + \varepsilon_{i,t}$$

Where $Fundreturn_{I,\ t-1}$ is the return (either gross or net monthly fund returns) of fund i in month t-1, α is a constant, $LOGTNA_{i,\ t-1}$ is the measure of fund size and $X_{i,t-1}$ is a set of control variables that includes LOGFAMSIZE, TURNOVER, AGE and EXPRATIO. $\varepsilon_{i,t}$ is a generic error term that is uncorrelated with all other dependent variables. $\varepsilon_{i,t}$ again stands for a generic error term that is uncorrelated with all other independent variables. β is the coefficient of interest in this regression as it captures the relationship between size and fund performance when controlled for a set of fund characteristics. γ stands for a vector of loadings on the control variables. I then take the estimates from these monthly regressions and take their time series means and standard deviations to form the overall estimates of effects of fund characteristics on performance.

Because this dataset is a set of panel data, I also estimate the above regression with monthly fixed effects and with firm fixed effects. This way both the cross-sectional as well as the time-series variation in the coefficients can be estimated and taken into account. Another point of interest is the standard error of the estimation. I use a robust clustered standard error at firm level to take clustering into account.

After this estimation I continue with a more detailed look at the relation between fund size and fund performance in family funds. the regression is specified as follows:

$$(5)\ Fundreturn_{i,t} = \alpha_i + \beta_{i,1} LOGTNA_{i,t-1} + \gamma X_{i,t-1} + \varepsilon_{i,t}$$

The coefficients of interest are the estimates for β_1 and β_2. For the hypothesis to be true both of the coefficients have to be positive and significant. It would mean that larger funds within fund families perform better than smaller funds within fund families.

The last part of this paper focusses on finding an optimal fund size and it first tries to do this by looking at both the returns and the cost side of mutual funds. To find if there is an optimal size for a mutual fund we examine the returns per decile first and then look at the regression results for these deciles. The optimal size should, firstly, have the highest average return or at least be close to the highest and, secondly, the coefficient for the total net assets in the regression analyses should show either a parabolic or a decreasing value. If the coefficient shows a parabolic shape it is clear to see that the optimal size is at the top of this parabola. If there is a decreasing value for the coefficient, the speed at which the coefficient decreases from decile to decile becomes of interest in finding the optimal size with regards to the return side of mutual funds.

The second part of this hypothesis looks at the estimated costs of a fund and if there is a pattern in these costs. The analysis is similar to the analyses for the return side as described above. The average expenses and management fees are estimated for each decile and checked for any patterns. This in combination with the coefficients estimated in hypothesis 1 will lead to conclusions about an optimal size with regards to the cost side of mutual funds.

VI. Results

Hypothesis 1

Table 8 reports the results from the regression specified in formula (4) using the Beta-Adjusted return with both month and firm fixed effects. As we can see from the estimated coefficients that size (LOGTNA) does seem to influence performance since the coefficients are all significant at a 5% or even at a 1% significance level. The coefficients are all negative which would mean that a positive shock to the size of a fund would lead to a decrease in fund return for that fund. This last finding is similar to that of Chen et. Al. (2004).

Table 8 Regression of Beta-Adjusted Fund Peformance on Lagged Fund Size

	Gross Fund Return			Net Fund Return		
	No fixed Beta-Adjusted	Month Fixed Beta-Adjusted	Month and firm fixed effects	No fixed Beta-Adjusted	Month Fixed Beta-Adjusted	Month and firm fixed effects
INTERCEPT	0,001	0,015	0,03	0,001	0,015	0,03
	(1,09)	(4,10)**	(6,12)**	(1,05)	(4,14)**	(6,16)**
$LOGTNA_{i,t-1}$	-0,0008	-0,0008	-0,0038	-0,0008	-0,00075	-0,0038
	(3,24)**	(3,03)**	(4,58)**	(3,23)**	(3,03)**	(4,58)**
$LOGFAMSIZE_{i,t-1}$	0,00007	0,00007	-0,00044	0,00007	0,00007	0,00044
	(0,36)	(0,4)	(0,48)	(0,36)	(0,4)	(0,48)
$TURNOVER_{i,t-1}$	-0,00022	-0,00021	-0,00027	-0,00022	-0,00021	-0,00027
	(3,96)**	(1,52)	(2,39)*	(3,98)**	(1,52)	(2,39)*
$AGE_{i,t-1}$	0,00012	0,00004	Omitted	0,00012	0,000038	Omitted
	(2,44)*	(2,75)**		(2,44)*	(2,75)**	
$EXPRATIO_{i,t-1}$	-0,093	-0,084	-0,08	-0,093	-0,084	(0,08)
	(1,95)	(1,9)	(1,11)	(1,95)	(1,9)	(1,11)
$MGMTFEE_{i,t-1}$	-0,00003	-0,00008	0,00091	-0,00003	-0,00008	0,00091
	(0,03)	(0,09)	(0,47)	(0,03)	(0,09)	(0,47)
$LAGFUNDRET_{i,t-1}$	0,0033	0,0087	0,01	0,0032	0,0087	0,01
	(2,72)**	(4,69)**	(3,53)**	(2,67)**	(4,69)**	(3,53)**
R^2	0,011	0,058	0,057	0,011	0,059	0,057
N	13,982	13,982	13,040	13,982	13,982	14,033

Table 8 estimated coefficients for regressions specified in formula 4 for the Beta-Adjusted return model. Shown above are the regressions with no fixed effects, with monthly fixed effects and with both month and firm fixed effects for both gross and net fund returns. The lagage variable is omitted in the month and firm fixed effects model as this variable does not provide any extra explanatory power to the fund returns. The estimated coefficients for the other models and for no fixed, monthly fixed and both month and firm fixed effects can be found in Appendix C at the end of this paper.

Other interesting results are the fact that the size of the family a fund is part of does not seem to be of influence to the performance of a fund. None of the estimated coefficients are statistically significant at both a 1% and 5% significance level. The sign of these coefficients is negative though. This would suggest that an increase in the size of a fund family leads to a decrease in fund performance. This is not consistent with earlier findings by for example Chen et. al. (2004).

Looking at the other estimated coefficients we see that the age of a fund is omitted from the regression that uses both month and firm fixed effects. This is done because the variable does not provide any extra explanatory power for fund returns. The turnover of a fund is a significant variable at a 5% significance level in each model. The sign is negative which means that a positive shock to the turnover of a fund leads to a decrease in fund return. Implications of this negative sign would suggest that active trading, which leads to a higher turnover for a fund, is not profitable for mutual funds. An interesting result since most active funds promote themselves for being able to hand pick the best stocks on the market. Following a passive strategy and not changing your portfolio much seems to be better than changing investments according to these coefficients.

The Expense ratio variable shows a negative sign which is what you would expect for a variable that approximates the cost side of mutual funds. An increase in the expense ratio of a fund should lead to a lower return and the coefficients seem to follow this same logic. One thing to note on these coefficients is the fact that none of them are significant at either 1% or 5% significance. The management fee variable shows however a positive sign which means that when the management fees increase, the returns also increase. This is not what I expected for this variable since the management fees are a cost to a mutual fund. The coefficients however are, once again, not significant at any level. It might be the case that a higher management fee indicates a better manager which is able to outperform managers more than he gets rewarded for it. Meaning that his relative salary is lower than that of a similar fund with lower performance.

The last variable used in the regression is the fund return of the previous period. This variable is significant for both significance levels and the coefficients are similar to results found by Cheng et al. (2004). The sign of the coefficients is positive which would mean that a positive return in the previous period would most likely lead to a positive return in the present period. A negative return in the previous period is likely to lead to a negative return in the present period. This effect is similar to the momentum effect as observed by Grinblatt et. Al. (1995). This effect says past winners continue on winning in the present and, at least in the near future, future and past losers continue on losing.

Looking at the differences between the different fixed effects regressions we can see that from the monthly fixed effects coefficients, as a firm grows over time the coefficients do not change a lot. The size effects seem to stay close to the same values as the regression without any fixed effects and the other return side variables, turnover and family size, stay the same as well. The size effect seems to be a little bit higher than with no fixed effects and therefore the size effects seems to be present in a time series setting. To also see if there is a size effect in a cross-sectional setting we have to look at the coefficients from the regression with both month and firm fixed effects. These coefficients will show us the size effect across large and small funds in the same month. Again, the size effect is a little larger than in the previous regressions. This shows us that in the same month, there is a difference between the performance of large and small funds which can be explained through the size of those funds. The only other change in coefficients can be seen in the Lagged fund return, this would mean that as a firm grows over time the past fund performance starts playing a more important role and that in the same month the difference between large and small funds can also be explained in a significant way through the past performance of those funds. The above analysis shows us that the size effect is observable in both a time-series as in a cross-sectional setting and that the magnitude of this effects seems to increase in a cross-sectional setting in comparison with a time-series setting.

Based on the above analyses of the estimated coefficients I accept the first hypothesis. This means that there is a negative relationship between the size of a mutual fund and its performance

The results from table 8 show that there is not a lot of difference in the estimated coefficients or their significance across the different models or between gross and net fund returns. To see if there are really decreasing returns to scale and if this relation is linear or not, I estimated the above regression for each decile using the CAPM model with net fund returns. I use the CAPM model as this is one of the most intuitive and easiest models to use whilst still being very applicable even in today's economic circumstances.

The result of this decile split can be found in table 9 below. The coefficients for the total net assets do not show a clear relation across deciles. There does not seem to be a clear linear relation between size and return. There are however more positive coefficients for the smaller fund deciles. Deciles 6 through 9 all show a negative sign and then decile 10 shows a positive coefficient again. It might be the case that the optimal size of a fund lies somewhere in between decile 4 and decile 5 and that after that point there are decreasing return to scale. But the positive coefficient for decile 10 show that it might be profitable to be one of the biggest funds in order to gain an abnormal return. This result is similar to the average returns per decile as shown in table 3, where the net return is increasing from decile 4 to decile 6 and starts decreasing again for the next deciles.

The estimated coefficients for the broken down regression formulated as in formula 4 do not lead to a different conclusion with regards to the acceptance of the first hypothesis. The first hypothesis is still accepted which means that there is a negative relationship between the size of a mutual fund and its performance.

Table 9 Regression of Fund Peformance on Lagged Fund Size using CAPM and net fund returns per decile with month and firm fixed effects

	CAPM model estimated with Net Fund Returns									
Deciles	1	2	3	4	5	6	7	8	9	10
INTERCEPT	0.141	0.042	0.008	0.038	0.040	0.038	-0.016	0.090	0.062	0.064
	(1.98)	(1.90)	(0.39)	(1.67)	(1.57)	(1.31)	(0.40)	(2.80)**	(1.67)	(2.57)*
$LOGTNA_{i,t-1}$	-0.001	0.002	-0.004	-0.001	0.000	-0.005	-0.003	-0.016	-0.016	0.004
	(0.07)	(0.24)	(0.48)	(0.10)	(0.00)	(0.70)	(0.41)	(2.33)*	(2.46)*	(1.08)
$LOGFAMSIZE_{i,t-1}$	0.001	-0.001	0.004	-0.003	-0.005	0.000	0.004	0.004	0.006	-0.008
	(0.04)	(0.31)	(0.88)	(0.61)	(1.61)	(0.06)	(1.08)	(1.06)	(1.48)	(2.61)**
$TURNOVER_{i,t-1}$	-0.001	-0.000	-0.001	-0.001	-0.000	-0.001	-0.000	0.000	0.001	0.001
	(0.72)	(2.68)**	(0.86)	(0.63)	(0.30)	(0.48)	(0.04)	(0.92)	(0.42)	(0.23)
$EXPRATIO_{i,t-1}$	-0.068	-0.795	-0.281	0.091	-0.115	-0.229	1.746	-1.550	0.149	1.201
	(1.30)	(1.39)	(0.60)	(0.16)	(0.24)	(0.29)	(1.48)	(1.37)	(0.10)	(0.60)
$MGMTFEE_{i,t-1}$	0.003	-0.005	0.002	-0.013	-0.007	-0.000	-0.001	0.009	-0.008	-0.017
	(1.30)	(0.69)	(0.46)	(2.85)**	(1.49)	(0.01)	(0.10)	(0.70)	(0.39)	(0.67)
$LAGFUNDRET_{i,t-1}$	0.036	0.003	0.008	0.010	0.011	0.006	0.001	0.024	0.007	0.014
	(2.39)*	(0.50)	(1.93)	(2.94)**	(1.67)	(0.92)	(0.18)	(4.99)**	(1.55)	(2.67)**
R^2	0.26	0.13	0.11	0.12	0.09	0.09	0.10	0.13	0.15	0.18
N	583	910	1,325	1,806	1,807	1,813	1,642	1,515	1,304	1,328

Table 12 estimated coefficients for the regression as specified in formula 4 for each size decile. The lagage variable is not included as this variable does not add extra explanatory power in explaining the fund returns. The regression uses both month and firm fixed effects. The monthly dummies are not shown here to keep the table easy to read. The regression estimates for the normal regression and the regression with only firm fixed effects can be found in appendix C at the end of this paper

Hypothesis 2

For the second hypothesis I only keep those firms that are part of a fund family and I include fund family fixed effects to investigate if within families there are decreasing returns to scale. The estimated coefficients can be found in table 10 below:

Table 10 Fund family fixed effects regression estimated to check for decreasing returns to scale within families

	Gross Fund Return				Net Fund Return			
	Market-Ad	Beta-Adj	3-Factor	4-Factor	Market-Adj	Beta-Adj	3-Factor	4-Factor
INTERCEPT	0,026	0,022	0,027	0,028	0,0176	0,0219	0,019	0,02
	(4,53)**	(3,86)**	(4,88)**	(5,09)**	(3,13)**	(3,86)**	(3,50)**	(3,72)**
$LOGTNA_{i,t-1}$	-0,0037	-0,0039	-0,004	-0,0038	-0,0036	-0,0039	-0,0039	-0,0038
	(3,60)**	(4,19)**	(3,92)**	(3,79)**	(3,84)**	(4,18)**	(4,23)**	(4,08)**
$LOGFAMSIZE_{i,t-1}$	-0,0012	-0,0011	-0,0008	-0,0009	-0,001	-0,0011	-0,00064	-0,00068
	(1,04)	(1,03)	(0,72)	(0,75)	(0,93)	(1,03)	(0,61)	(0,65)
$TURNOVER_{i,t-1}$	-0,0004	-0,0002	-0,0004	-0,0004	-0,00018	-0,00016	-0,00016	-0,00018
	(5,02)**	(3,40)**	(5,69)**	(5,79)**	(3,50)**	(3,39)**	(3,91)**	(4,34)**
$EXPRATIO_{i,t-1}$	0,237	0,255	0,183	0,17	0,277	0,254	0,222	0,21
	(1,04)	(1,1)	(0,83)	(0,78)	(1,19)	(1,1)	(0,99)	(0,95)
$MGMTFEE_{i,t-1}$	-0,0039	-0,0047	-0,0038	-0,0036	-0,0047	-0,0047	-0,0046	-0,0045
	(1,3)	(1,68)	(1,29)	(1,22)	(1,69)	(1,69)	(1,67)	(1,62)
$LAGFUNDRET_{i,t-1}$	0,0011	0,0021	0,0015	0,0017	0,0013	0,0021	0,0018	0,002
	(0,58)	(1,15)	(0,83)	(0,93)	(0,69)	(1,12)	(0,98)	(1,08)
R^2	0,016	0,017	0,016	0,015	0,014	0,016	0,014	0,014
N	6,084	6,441	6,084	6,084	6,441	6,441	6,441	6,441

Table 10 estimated coefficients for regression as specified in formula 4. The lagage variable is not included as it did not provide any extra explanatory power in explaining the performance of a fund. This regression includes fund family fixed effects and monthly interaction effects. The interaction dummy coefficients are not shown here to keep the table easy to read

From the estimated coefficients we can see a couple of implications for the relation between the size of a fund within a family and the relation with fund return. Firstly, when looking at the coefficients for the size of a fund the estimates are all highly significant and show a negative sign. This suggest that when the size of a fund within a family increases it return goes down. This would suggest there are decreasing returns to scale within fund families.

Secondly, the size of the family itself influences the return of the funds within in a negative way as well. A larger fund family tends to perform worse than a smaller fund family, although these coefficients show no significance at either 5% or 1% significance level. Thirdly, the turnover of a fund still is very significant in predicting the return of a fund, like seen before the higher the turnover the lower the return seems to be. The expense ratio shows a positive coefficient which is not what you would expect from a cost variable but none of the coefficients are significant. The coefficients of the management fee show the expected negative signs but again are not significant.

The last interesting observation from these coefficients is the funds return from the previous period. This variable was significant in the first set of estimated regressions but when looking at this regression with fund family fixed effects it loses this significance. The sign of the variable has not changed which means that the interpretation still remains the same and that the momentum effect still plays somewhat of a role.

The above analyses of the estimated coefficients for the regression as specified in formula 4 and with family fixed effects, leads to the rejection of the second hypothesis. This means that there is not enough evidence to support a positive relationship between the size of a fund within a family and its return. There is not enough evidence to suggest that a large firm within a fund family has a higher return than a smaller fund within that same family.

Hypothesis 3

To find an optimal size of a fund we have to look at all of the above estimated coefficients and descriptive statistics from the beginning of this paper.

On the return side of mutual funds, we can see in table 3 that for gross returns, the average return per decile is decreasing when moving up in deciles. This is what was expected at the start of the paper as decreasing returns to scale for mutual funds has been found before (Chen, Hong, Hiang, & Kubik, 2004). But when looking at the net returns in the same table we do not observe such a clear pattern as with the gross returns. Average returns seem to be increasing for the first 6 deciles and decrease again after that size decile. This would suggest that the optimal size must be somewhere in or around that decile. If we take the regression results from table 9 into account, we saw that the relation between size and return again is not as linear as expected. For the first 4 deciles size seems to be a negative influence on return, then for deciles 5 and 6 the relation goes briefly positive and the deciles after that are negative again with the exception of the biggest size decile. Again the optimal size seems to be around the 5th or 6th size decile. The regression coefficients from table 9, however, are not significant at either 5% or 1% significance.

On the cost side of mutual funds, we can see from table 11 below that for the expense ratio there is a clear decreasing pattern across deciles. The decrease in the expense ratio does slow down the higher a fund goes in the decile portfolios. When looking at the management costs across deciles there is not as clear as a pattern as with the expense ratio. For decile 1 through 4 there is a decrease in management cost but after that the costs start increasing again up until the largest size decile which has lower management costs than the deciles that are just smaller than that. For the management costs the optimal size would be somewhere in decile 4 because after that point the management cost start to increase again. It is not the decile with the lowest management cost but it is very close to the same number as decile 10. For the expense ratio it is not so clear where the optimum size is at. The lowest ratio is at decile 10 so this would be the point at which a fund has the lowest ratio, but looking at the marginal decreases the decrease from going from decile 1 to decile 2 leads to a decrease of 1,19%. This marginal decrease decreases itself over the size deciles as well with the lowest decrease being the step from decile 6 to decile 7, 0,07%.

Table 11 Average Expense ratios and Management Costs per decile and for the total sample

	Decile 1	Decile 2	decile 3	Decile4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Total
Expenses	3,20%	2,01%	1,88%	1,64%	1,55%	1,45%	1,38%	1,28%	1,19%	1,02%	1,55%
	0,57%	0,06%	0,05%	0,03%	0,03%	0,02%	0,02%	0,02%	0,02%	0,03%	0,04%
Management Costs	0,91	0,85	0,86	0,79	0,83	0,86	0,87	0,86	0,86	0,78	0,84
($ million)	0,12	0,03	0,03	0,02	0,02	0,02	0,01	0,02	0,02	0,02	0,01

Table 11 Average Expense Ratios and Management Costs per decile and for the total sample

If we take the descriptive statistics as described above in combination with the estimated coefficients from the regression analyses, we can find the optimum size with regards to cost efficiency for mutual funds. From table 9 we can see that the coefficients for the expense ratio are mostly negative which is as expected but for the highest two deciles the coefficient is positive. Because a positive coefficient for the expense ratio does not make a lot of sense since one would expect that an increase in the expense ratio would lead to a decrease in the return of a fund, I only look at the negative values and decile 5 is the decile with the lowest coefficient. An increase in the expense ratio in this decile would lead to the smallest decrease in fund return and would therefore be optimal. The management cost shows negative values for all but the first and smallest decile. The lowest value is decile 6, an increase in the management costs would lead to the smallest decrease in fund return and would therefore be optimal.

The third hypothesis was a search for an optimal size for a mutual fund. In this hypothesis I have looked at both the return as well as the cost side of mutual funds to see if they both show similar results for an optimal size decile. Looking at the above analyses we can see that the optimal size for both the return as well as the cost side would be somewhere in and or in between the 5^{th} and the 6^{th} size decile. To get a clearer idea of how much net assets under management that would be, we have to look at the descriptive statistics for each decile found in table 4. From this table we can see that decile 5 has a minimum value for net assets under management of $10,2 million and the maximum of decile 6 is $58,80 million. This range is very large and could be brought down with a more detailed review of funds within that range and their estimated coefficients, but this goes beyond the scoop of this paper.

VII. Conclusion

In this paper I have tried to investigate the true relation between the size of mutual funds and their performance. In the light of the recent growth spurt, which is still ongoing, of mutual funds and their increasing popularity amongst normal people this relation is an important establishment for every investor on the market.

The main question of this paper was as follows:

Does fund size erode mutual fund performance?

To find a clear and profound answer to this question, I split the question into three different hypotheses which each looked at a different aspect of the relationship between size and performance. The analyses of the first hypothesis lead to the acceptance of the hypothesis and confirmed the results of previous research that there does seem to be a negative relationship between the size of a fund and its performance. Even a more detailed analyses of this result when the regression was split into 10 size deciles, showed enough evidence to suggest that there is such a negative relationship. This last analyses however made clear that the decreasing returns to scale are not of a linear form. There does seem decently large differences between deciles, this information would later be used in answering the third hypothesis.

The second hypothesis looked at decreasing returns to scale within fund families. This hypothesis was researched by using a fixed effect regression to only get the coefficients within fund families. The estimated coefficients and the complimentary analyses lead to the rejection of the second hypothesis. This implicates that there is not enough evidence to support a negative relationship between size and fund performance within fund families.

The third hypothesis looked at finding an optimal size for a mutual fund. The optimal size for this sample of mutual funds lies somewhere in the 5^{th} or 6^{th} size decile. This translates into a range from $18,6 million to $58,8 million net assets under management. With the chosen methods in this paper and the resources at hand, this was the smallest range I could find.

With regards to answering the main question of this paper I think it is safe to say that fund size does in fact erode mutual fund performance. the results show negative and significant coefficients for fund size and increasing values for the coefficients when moving up in the size deciles. Even within families, fund size seems to be eroding mutual fund performance. If fund size would not erode mutual fund performance the optimal fund size would be somewhere in decile 10, but this optimal size lies somewhere in the 5^{th} or 6^{th} decile.

The implications of this research reach down to individual investor level. The results implicate that investing in bigger funds is not per se better. Picking the right mutual fund to get the highest return possible on your investment does not seem to be as clear cut as that. The research implicates also that for fund families it might be best to split their funds into smaller more specialized funds. a fund family could for example consist of a mutual fund for each equity market or for each continent. The size of a fund within a family would be kept small in this case and this would be more beneficial as we have seen from the estimated coefficients that there is not enough evidence to support a positive relationship between size of a fund and performance within a fund family. The last implication would be for fund managers. They should not be concerned with have the largest pile of money under their management but they should split their assets under management whenever they hit a certain threshold in order to maintain the highest return possible on their investments.

Shortcomings of this paper are the fact that the dataset is not as complete as I wanted it to be. Some of the values for the assets under management and the reported returns are missing for the time period. This also goes for the variables used on the cost side of mutual funds. Another shortcoming is the fact that the cost side of a mutual fund is only measured by using the generic expense ratio and the costs for the management. Splitting this expense ratio into different expenses such as administrative and marketing for example would make for a more detailed analyses and might lead to different results and conclusions. Another shortcoming regards the chosen time period for the analyses. Within the chosen time period lies a very big financial crisis which hit the financial markets very hard and made a big impact on fund performance. this might be the reason that the factor loadings show an opposite pattern than what I expected before starting this analyses.

For future research I would suggest trying to find more, complete, observations. This would make the entire analyses more profound as a whole. Another suggestion if the same time period is used is to split the sample in two to count for the fact of the financial crises being in the sample. The financial crisis had a huge impact on financial markets and might have led to a survival bias in the sample which might have influenced the results and the accompanied conclusions. With regards to finding an optimal size for a mutual fund a more detailed split of the used variables on the cost side would possible lead to better and or different results. Also extending the used variables with more and or different costs and or return measuring variables would increase the effectiveness of the research.

VIII. Appendix A: Hypothesis 1 tables

Table A1 Regression of Fund Peformance on Lagged Fund Size without fixed effects

	Gross Fund Return				Net Fund Return			
	Market-Adj	Beta-Adj	3-Factor	4-Factor	Market-Adj	Beta-Adj	3-Factor	4-Factor
INTERCEPT	-0,001	0,001	0,002	0,003	-0,002	0,001	0,001	0,002
	(0,38)	(1,09)	(1,19)	(1,82)	(1,6)	(1,05)	(0,77)	(1,7)
$LOGTNA_{i,t-1}$	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001
	(3,22)**	(3,24)**	(3,88)**	(3,53)**	(2,13)*	(3,23)**	(2,83)**	(2,49)*
$LOGFAMSIZE_{i,t-1}$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	(0,51)	(0,36)	(0,57)	(0,6)	(0,23)	(0,36)	(0,28)	(0,35)
$TURNOVER_{i,t-1}$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	(4,70)**	(3,96)**	(4,56)**	(4,42)**	(3,62)**	(3,98)**	(3,30)**	(3,17)**
$AGE_{i,t-1}$	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	(0,98)	(2,44)*	(0,88)	(0,9)	(2,55)*	(2,44)*	(2,51)*	(2,60)**
$EXPRATIO_{i,t-1}$	0,129	-0,093	0,131	0,132	-0,095	-0,093	-0,093	-0,093
	(2,26)*	(1,95)	(2,31)*	(2,33)*	(1,96)	(1,95)	(1,97)*	(1,95)
$MGMTFEE_{i,t-1}$	0,002	0,0000	0,002	0,002	0,0000	0,0000	0,0000	0,0000
	(1,66)	(0,03)	(1,59)	(1,63)	0	(0,03)	(0,1)	(0,06)
$LAGFUNDRET_{i,t-1}$	0,002	0,003	0,003	0,003	0,003	0,003	0,003	0,004
	(1,61)	(2,72)**	(2,08)*	(2,22)*	(2,25)*	(2,67)**	(2,82)**	(2,99)**
R^2	0,02	0,01	0,02	0,02	0,01	0,01	0,01	0,01
N	12,994	13,982	12,994	12,994	13,982	13,982	13,982	13,982

Table A1 estimated coefficients for regressions specified in formula 4 for each of the used models. The regressions use neither month nor firm fixed effects.

Table A2 Regression of Fund Peformance on Lagged Fund Size with month fixed effects

	Gross Fund Return				Net Fund Return			
	Market-Adj	Beta-Adj	3-Factor	4-Factor	Market-Adj	Beta-Adj	3-Factor	4-Factor
INTERCEPT	0,017	0,015	0,011	0,012	0,013	0,015	0,007	0,008
	(4,34)**	(4,10)**	(2,74)**	(3,03)**	(3,62)**	(4,14)**	(1,88)	(2,20)*
$LOGTNA_{i,t-1}$	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001
	(3,25)**	(3,03)**	(3,90)**	(3,55)**	(2,32)*	(3,03)**	(3,00)**	(2,66)**
$LOGFAMSIZE_{i,t-1}$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	(0,57)	(0,4)	(0,56)	(0,61)	(0,32)	(0,4)	(0,29)	(0,38)
$TURNOVER_{i,t-1}$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	(2,96)**	(1,52)	(3,06)**	(2,92)**	(1,47)	(1,52)	(1,55)	(1,44)
$AGE_{i,t-1}$	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	(1,18)	(2,75)**	(1,18)	(1,2)	(2,77)**	(2,75)**	(2,75)**	(2,85)**
$EXPRATIO_{i,t-1}$	0,138	-0,084	0,139	0,14	-0,085	-0,084	-0,084	-0,084
	(2,57)*	(1,9)	(2,59)**	(2,59)**	(1,93)	(1,9)	(1,91)	(1,9)
$MGMTFEE_{i,t-1}$	0,002	0,0000	0,002	0,002	0,0000	0,0000	0,0000	0,0000
	(1,63)	(0,09)	(1,58)	(1,61)	(0,05)	(0,09)	(0,11)	(0,07)
$LAGFUNDRET_{i,t-1}$	0,008	0,009	0,008	0,008	0,009	0,009	0,009	0,009
	(3,80)**	(4,69)**	(3,83)**	(3,81)**	(4,69)**	(4,69)**	(4,72)**	(4,70)**
R^2	0,07	0,06	0,05	0,05	0,07	0,06	0,04	0,04
N	12,994	13,982	12,994	12,994	13,982	13,982	13,982	13,982

Table A2 estimated coefficients for regressions specified in formula 4 for each of the used models. The regressions use month fixed effects but no firm fixed effects. The monthly dummy variables are not shown here to keep the table easy to read.

Table A3 Regression of Fund Peformance on Lagged Fund Size with month and firm fixed effects

	Gross Fund Return				Net Fund Return			
	Market-Adj	Beta-Adj	3-Factor	4-Factor	Market-Adj	Beta-Adj	3-Factor	4-Factor
INTERCEPT	0,035	0,03	0,028	0,029	0,028	0,03	0,022	0,023
	(7,12)**	(6,12)**	(5,78)**	(6,04)**	(5,78)**	(6,16)**	(4,43)**	(4,70)**
$LOGTNA_{i,t-1}$	-0,0035	-0,0038	-0,0038	-0,0037	-0,0036	-0,0038	-0,0039	-0,0038
	(3,64)**	(4,58)**	(3,91)**	(3,81)**	(4,31)**	(4,58)**	(4,63)**	(4,51)**
$LOGFAMSIZE_{i,t-1}$	-0,0006	-0,0004	-0,0005	-0,0005	-0,0005	0,0004	-0,0004	-0,0004
	(0,57)	(0,48)	(0,5)	(0,5)	(0,57)	(0,48)	(0,47)	(0,48)
$TURNOVER_{i,t-1}$	-0,0005	-0,0003	-0,0005	-0,0005	-0,0003	-0,0003	-0,0003	-0,0003
	(4,09)**	(2,39)*	(4,17)**	(4,18)**	(2,41)*	(2,39)*	(2,41)*	(2,55)*
$EXPRATIO_{i,t-1}$	0,033	-0,080	0,034	0,034	-0,081	-0,080	-0,080	-0,079
	(0,42)	(1,11)	(0,42)	(0,43)	(1,12)	(1,11)	(1,11)	(1,1)
$MGMTFEE_{i,t-1}$	0,0028	0,00091	0,0028	0,0028	0,00095	0,00091	0.00092	0,00091
	(1,35)	(0,47)	(1,33)	(1,33)	(0,5)	(0,47)	(0,48)	(0,47)
$LAGFUNDRET_{i,t-1}$	0,009	0,0100	0,009	0,009	0,0100	0,0100	0,0100	0,0100
	(3,00)**	(3,53)**	(3,02)**	(3,01)**	(3,53)**	(3,53)**	(3,55)**	(3,54)**
R^2	0,064	0,057	0,04	0,039	0,066	0,057	0,039	0,038
N	13,04	13,04	13,04	13,04	14,033	14,033	14,033	14,033

Table A3 estimated coefficients for regression specified in formula 4 for each of the used models. The regressions use both month and firm fixed effects. The monthly dummy variables are not shown here to keep the table easy to read.

Table A4 Normal Regression of Fund Peformance on Lagged Fund Size using CAPM and net fund returns per decile

	CAPM model estimated with Net Fund Returns									
Deciles	1	2	3	4	5	6	7	8	9	10
INTERCEPT	0.001	0.005	-0.007	-0.009	-0.014	-0.012	-0.001	-0.020	-0.022	0.009
	(0.16)	(0.60)	(0.71)	(0.69)	(0.88)	(0.77)	(0.07)	(0.93)	(1.19)	(1.25)
$LOGTNA_{i,t-1}$	0.003	0.002	0.003	0.005	0.004	0.003	0.001	0.004	0.002	-0.002
	(0.82)	(0.36)	(0.73)	(1.05)	(0.74)	(0.79)	(0.28)	(0.93)	(0.59)	(1.73)
$LOGFAMSIZE_{i,t-1}$	-0.000	0.000	-0.000	0.000	0.000	-0.000	-0.000	0.000	0.001	0.001
	(0.44)	(0.09)	(0.06)	(0.37)	(0.70)	(0.25)	(0.20)	(0.18)	(1.03)	(0.65)
$TURNOVER_{i,t-1}$	-0.000	0.000	0.000	-0.000	-0.000	-0.000	-0.000	0.001	-0.002	-0.001
	(17.31)**	(1.31)	(0.34)	(0.58)	(0.66)	(0.35)	(1.09)	(2.70)**	(4.05)**	(1.22)
$AGE_{i,t-1}$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.80)	(1.80)	(2.46)*	(1.38)	(1.19)	(1.06)	(0.63)	(1.42)	(0.89)	(2.87)**
$EXPRATIO_{i,t-1}$	-0.062	-0.639	-0.192	-0.391	0.004	-0.389	-0.113	-0.012	1.138	-0.480
	(1.52)	(3.17)**	(1.16)	(2.09)*	(0.02)	(1.21)	(0.53)	(0.04)	(2.58)*	(1.76)
$MGMTFEE_{i,t-1}$	0.001	0.001	0.001	0.000	-0.002	0.004	-0.004	-0.005	-0.012	0.001
	(0.52)	(0.35)	(0.47)	(0.02)	(1.18)	(1.10)	(1.17)	(1.10)	(1.99)*	(0.40)
$LAGFUNDRET_{i,t-1}$	0.023	-0.005	0.005	0.004	0.000	0.001	-0.000	0.006	0.000	0.003
	(2.29)*	(1.14)	(1.86)	(2.06)*	(0.11)	(0.29)	(0.09)	(2.62)**	(0.07)	(1.23)
R^2	0.08	0.06	0.01	0.01	0.00	0.00	0.01	0.01	0.02	0.01
N	583	909	1,320	1,790	1,798	1,812	1,641	1,514	1,299	1,316

Table A4 estimated coefficients for the regression as specified in formula 4 for each size decile. The regression uses neither month nor firm fixed effects.

Table A5 Regression of fund performance on Lagged Fund Size using CAPM and net fund returns per decile with month fixed effects

	CAPM model estimated with Net Fund Returns									
Deciles	1	2	3	4	5	6	7	8	9	10
INTERCEPT	0.104	0.013	0.005	-0.007	-0.004	0.006	0.003	0.033	-0.034	0.026
	(1.66)	(0.71)	(0.37)	(0.45)	(0.22)	(0.34)	(0.14)	(1.34)	(1.40)	(1.77)
$LOGTNA_{i,t-1}$	0.000	0.003	0.004	0.006	0.004	0.002	0.002	0.001	0.003	-0.002
	(0.01)	(0.65)	(0.89)	(1.44)	(0.88)	(0.42)	(0.50)	(0.23)	(0.70)	(1.60)
$LOGFAMSIZE_{i,t-1}$	-0.000	0.000	0.000	0.000	0.001	-0.000	-0.000	-0.000	0.001	0.000
	(0.40)	(0.08)	(0.14)	(0.39)	(1.27)	(0.83)	(0.50)	(0.10)	(1.07)	(0.19)
$TURNOVER_{i,t-1}$	-0.000	0.000	-0.000	-0.000	-0.001	-0.000	-0.001	0.001	-0.002	-0.002
	(11.62)**	(1.63)	(0.39)	(1.09)	(0.96)	(0.34)	(1.55)	(2.03)*	(3.46)**	(1.42)
$AGE_{i,t-1}$	-0.000	0.000	0.000	0.000	0.000	-0.000	-0.000	0.000	-0.000	0.000
	(0.42)	(1.19)	(1.33)	(0.69)	(0.89)	(0.81)	(0.24)	(0.44)	(0.51)	(1.74)
$EXPRATIO_{i,t-1}$	-0.049	-0.627	-0.216	-0.319	0.007	-0.337	-0.088	-0.151	1.020	-0.488
	(1.44)	(3.03)**	(1.35)	(1.76)	(0.04)	(1.10)	(0.39)	(0.43)	(2.34)*	(1.60)
$MGMTFEE_{i,t-1}$	0.002	0.000	0.003	-0.000	-0.002	0.004	-0.003	-0.005	-0.012	0.000
	(1.04)	(0.10)	(1.18)	(0.22)	(1.21)	(0.96)	(1.14)	(0.99)	(2.29)*	(0.09)
$LAGFUNDRET_{i,t-1}$	0.035	-0.001	0.009	0.009	0.008	0.006	0.003	0.012	0.002	0.007
	(2.94)**	(0.14)	(2.71)**	(2.99)**	(1.90)	(1.51)	(0.87)	(2.99)**	(0.74)	(2.11)*
R^2	0.24	0.16	0.10	0.12	0.09	0.08	0.10	0.10	0.13	0.14
N	583	909	1,320	1,790	1,798	1,812	1,641	1,514	1,299	1,316

Table A5 estimated coefficients for the regression as specified in formula 4 for each size decile. The regression uses month fixed effects but no firm fixed effects. The monthly dummy variables are not shown here to keep the table easy to read

IX. Bibliography

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance, 52*(1), 57-82.

Center for Research in Security Prices. (sd). *Lipper Objective and Classification Codes*. Opgehaald van http://www.crsp.com/products/documentation/lipper-objective-and-classification-codes

Chen, J., Hong, H., Hiang, M., & Kubik, J. (2004, December). Does fund size erode mutual fund performance? The role of liquidity and organization. *The american economic review, 94*(5), 1276-1302. Opgeroepen op 2016

Cuthbertson, K., Nitzsche, D., & O'Sullivan, N. (2010, April 16). Mutual fund performance: Measurement and Evidence. *Financial Markets, Institutions & Instruments, 19*(2), 95-187. Opgeroepen op 2016, van http://onlinelibrary.wiley.com.eur.idm.oclc.org/doi/10.1111/j.1468-0416.2010.00156.x/abstract

Fama, E. F., & French, K. R. (1993). Common Risk Factors in the Return of Stocks and Bonds. *Journal of Financial Economics, 33*(1), 3-56.

Fortin, R., & Michelson, S. (1999, Feb). Fund Indexing Vs. Active Management: the results are... *Journal of Financial Planning, 12*(2), 74-81. Opgeroepen op 2016

Grinblatt, M. (1996, July). Another puzzle: the growth in actively managed mutual funds. *The journal of Finance, 51*(3), 783-810.

Grinblatt, M., & Titman, S. (1992, December). The persistence of mutual fund performance. *The journal of Finance, 47*(5), 1977-1984. Opgeroepen op 2016, van http://www.jstor.org.eur.idm.oclc.org/stable/2329005?seq=1#page_scan_tab_contents

Grinblatt, M., Titman, S., & Wermers, R. (1995). Momentum investment strategies, portfolio performance and herding: A study of mutual fund behavior. *The American Economic Review*(85), 1088-1105. Opgeroepen op 2016, van http://search.proquest.com.eur.idm.oclc.org/docview/233017216?accountid=13598

Indro, D., Jiang, C., Hu, M., & Lee, W. (1999, May). Mutual fund performance: does size matter? *Financial analysis journal, 55*(3), 74-87. Opgeroepen op 2016

Investment Company Institute. (2015). *Investment Company Fact Book 2015.* Investment Company Institute, Washington. Opgeroepen op 2016, van https://www.ici.org/pdf/2015_factbook.pdf

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance, 48*(1), 65-91.

Lowenstein, R. (1997, November 20). Frightened funds: is there a master in the house. *Wall Street Journal*, C1.

Perold, A., & Salomon, R. S. (1991). The right amount of assets under management. *Financial Analysts Journal, 3*(47), 31-39.

Perold, A., & Salomon, R. S. (1991). The right amount of assets under management. *Financial Analysis Journal, 47*(3), 31-39. Opgeroepen op 2016

Reuter, J., & Zitzewitz, E. (2010). How much does size erode mutual fund performance? A regression discontinuity approach. *NBER working paper series*. Cambridge. Opgeroepen op 2016

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance, 19*(3), 425-442. Opgeroepen op 2016