Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 14 (2007) 662–693
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Rating mutual funds: Construction and information content of an investor-cost based rating of Danish mutual funds ☆

Ken L. Bechmann [1], Jesper Rangvid *

Department of Finance, Copenhagen Business School, Solbjerg Plads 3, DK-2000 Frederiksberg, Denmark

Received 9 March 2006; received in revised form 17 February 2007; accepted 6 March 2007
Available online 17 May 2007

Abstract

We develop a new rating of mutual funds: the atpRating. The atpRating assigns crowns to each individual mutual fund based upon the costs an investor pays when investing in the fund in relation to what it would cost to invest in the fund's peers. Within each investment category, the rating assigns five crowns to funds with the lowest costs and one crown to funds with the highest costs.

We investigate the ability of the atpRating to predict the future performance of a fund. We find that an investor who has invested in the funds with the lowest costs within an investment category would have obtained a risk-adjusted excess return that is approximately 3–4 percentage points higher per annum than if the funds with the highest costs had been invested in.

We compare the atpRating with the Morningstar Rating. We show that one reason why the atpRating and the Morningstar Rating contain different information is that the returns Morningstar uses as inputs when rating funds are highly volatile whereas the costs the atpRating uses as inputs when rating funds are highly persistent. In other words, a fund that has low costs one year will most likely also have low costs the following year, whereas the return of a fund in a certain year generally contains only little information about the future return that the fund will generate.

☆ We thank ATP for its financial support. The mutual fund data that we use have kindly been supplied by Morningstar through ATP. We thank two anonymous referees, Magnus Dahlquist, Martin Gruber, Martin Holmén, Michael Møller, Johannes Raaballe, market participants from the Danish mutual fund industry, in particular Mads Gosvig, Bjarne Graven Larsen, Carsten Koch, and Jesper Kirstein, as well as participants at The Multinational Finance Conference and at seminars at the Aarhus School of Business, Copenhagen Business School, the Danish Central Bank, Stockholm Institute for Financial Research, and the Danish Society of Financial Analysts for useful comments. Morningstar and atpRating are trademarks.

* Corresponding author. Tel.: +45 3815 3784; fax: +45 3815 3615.
E-mail addresses: kb.fi@cbs.dk (K.L. Bechmann), jr.fi@cbs.dk (J. Rangvid).
[1] Tel.: +45 3815 2952.

doi:10.1016/j.jempfin.2007.03.001

Finally, we have information on the investments in different mutual funds made by a small subgroup of investors known to have been exposed to both the atpRating and the Morningstar Rating. We find that investors have clear preferences for funds rated high by both the atpRating and the Morningstar Rating.

JEL classification: G11; G23
Keywords: Mutual fund ratings; Mutual fund costs and performances; Performance measurement; Investor choices

1. Introduction

In 2003, there were 215 stocks listed on the Copenhagen Stock Exchange. At the same time there were 363 Danish mutual funds available to investors. In the United States, there are likewise more funds than there are listed stocks. In order to help investors monitor and choose between mutual funds, rating agencies have been created. Mutual fund ratings, such as Morningstar's, summarize useful information for the investor; For instance, Blake and Morey (2000) show that Morningstar Ratings can help investors avoid the worst performing funds. Mutual fund ratings are also potentially relevant to the mutual funds themselves; Del Guercio and Tkac (2001) show, for instance, that high-rated funds receive abnormally high inflows.

In this paper, we introduce and analyze a new rating of mutual funds. To do so, we first develop an indicator of the size of costs in Danish mutual funds. We then use the cost indicator of each individual fund to rate the funds according to the size of their cost indicators. The mutual fund rating we construct has been given the name atpRating.[2] The atpRating assigns crowns to each individual mutual fund, where the number of crowns a fund receives is determined by the costs of the fund compared to the costs of the fund's peers; a high-cost fund receives one crown and a low-cost fund receives five crowns. We test the predictive power of the atpRating, explain and analyze how it differentiates itself from Morningstar's rating, and study whether individual private investors use the information contained in the new atpRating.

There are two main reasons why we construct a new cost-based rating of mutual funds. First, Barber, Odean and Zheng (2005) hypothesize that "expenses that remain out of sight are likely to remain out of mind". In other words, the way information about costs is conveyed to investors affects how investors perceive and learn about the costs of mutual fund investing. The underlying hypothesis for developing a new cost indicator is thus that it might be difficult for ordinary investors to absorb and correctly understand the vast amount of information about the costs of mutual fund investing. In particular, several different types of costs have to be identified, weighted together and compared to funds with a similar investment strategy. Ordinary investors need to correctly understand and use the several kinds of costs associated with mutual fund

[2] ATP (*Arbejdsmarkedets Tillægspension*) is a Danish pension scheme to which all Danes pay mandatory contributions. ATP is one of the largest pension managers in Europe, managing assets worth more than DKK 300 billion, or approximately USD 53 billion. Until January 1, 2005, Danes could not influence the portfolio composition of their pension savings in ATP. As of January 1, 2005, however, all Danish individuals have been allowed to allocate a part of their pension savings (the Special Pension Savings Scheme; *Den Særlige Pensionsordning*) into different mutual funds (the aggregate value of the Special Pension Savings Scheme is DKK 43 billion, or approximately USD 8 billion). To facilitate the investment decisions of individual Danes, ATP initiated the development of the atpRating. It is the development of this rating that we describe and analyze in this paper.

investments because the costs of mutual fund investments influence the return to the investor (Elton et al., 1993; Malkiel, 1995; Gruber, 1996; Carhart, 1997; Chalmers et al., 1999).[3] To help investors better understand the costs of mutual fund investing, a simple transparent cost rating could be beneficial.

The second reason for developing a cost-based rating is due to our empirical findings showing that costs are highly persistent (if, for instance, a fund has low costs one year it is also likely to have low costs the following year), whereas returns are highly volatile (if, for instance, a fund has delivered a high return this year, the likelihood of it also delivering a high return next year is not, on average, high; see also Carhart, 1997). A rating based on the most recent costs thus has the potential to contain reliable information about future costs and thereby perhaps about future returns to the investor.

After developing the indicator of the size of mutual fund costs as well as rating the funds, we test the predictive power of the cost rating for future risk-adjusted returns using data from the Danish mutual fund market for the period 1994 to 2003. We find that the rating contains information about the future long-run performance of the funds and, in particular, that an investment in a low-cost fund, on average, is associated with a higher risk-adjusted return than is an investment in a high-cost fund if the investor holds the fund over eight- to ten-years. Over shorter investment horizons (three to five years), we find that costs do not appear to predict performance.

It should be mentioned explicitly that we test whether investing in a fund with, for instance, low cost in period t on average delivers higher returns in period $t+1$, $t+2$, and so forth, i.e. whether there is a relation between the size of a fund's costs in the *current* period and the *future* performance of the fund. Indeed, by focusing on the predictive power of current costs, we differentiate ourselves from Elton et al. (1993), Malkiel (1995), Carhart (1997), and Chalmers et al. (1999) who test whether there is an *in-sample* relation between expenses and returns, and report this relation to be generally significantly negative. Gruber (1996) also investigates whether there is a relation between costs today and future performances of a fund. Apart from our focus on a market other than the US market, which Gruber analyses, the contribution of this paper is that we develop a performance metric based on mutual fund costs and use it to examine performance predictability spanning both short horizons (three years) and long horizons (ten years). In addition, we study the whole universe of Danish mutual funds (active and index funds, domestic and foreign funds, equity and bond funds, and so forth) whereas Gruber looks at domestic equity funds.

Today, there are two ratings of Danish mutual funds: the atpRating and the Morningstar Rating. We thus compare the atpRatings with the Morningstar Ratings and find that the two ratings contain different information about mutual funds. We also discuss the possible reasons underlying these differences, as well as their implications for the predictive power of the two ratings.

Finally, evidence is provided concerning the importance of the atpRating for the inflow of cash to mutual funds. In particular, we track the size of inflows to mutual funds when savings in the Special Pension Savings Scheme[4] have been allocated between different individual mutual funds.

[3] Choi et al. (2006) examine and confirm in their experimental study that the way costs about mutual fund investments are brought to the attention of investors influences investors' perceptions of the attractiveness of a given fund. In particular, Choi et al. show that investors "overwhelmingly fail to minimize their index fund fees" unless investors are made clearly aware of the size of the fees and the effect of fees on otherwise similar index funds' performances.

[4] See footnote 1.

The data used here have a unique feature in that we know that when investors allocated their pension savings to different mutual funds, they were exposed to *both* the Morningstar Rating as well as the atpRating because of the information provided by the Internet platform that must be used when allocating pension savings between the different mutual funds. This feature of the data allows us to present summary statistics on the inflow to funds rated by Morningstar and funds rated by the atpRating. We find that investors clearly favor funds rated high by either the Morningstar Rating or the atpRating. These findings add to those of Del Guercio and Tkac (2001), who show that Morningstar Ratings strongly affect the inflows to US mutual funds.

Before commencing, the reasons why the *Danish* mutual fund market is studied in our analysis of the cost-based mutual fund rating should also be mentioned. First and foremost, the atpRating was developed for use by ordinary Danes when making investment decisions regarding their savings in the Special Pension Savings Scheme. Second, as just explained, data on these investment decisions allow us to investigate the extent to which investors rely on both the Morningstar Rating and the atpRating. Finally, by analyzing Danish mutual funds, we provide non-US, and in this sense out-of-sample, evidence on the issue of whether high-cost funds generate higher returns than low-cost funds, thereby contributing to the work of e.g. Blake et al. (1993), Gruber (1996), Carhart (1997), and Wermers (2000), who study US mutual funds.[5]

After these introductory remarks, the remaining part of the paper is structured as follows. In the next section, the Danish mutual fund market is described briefly. In Section 3, the cost structure of Danish mutual funds is explained and documented. Section 4 describes the construction of the atpRating, and in Section 5, whether or not the atpRating contains useful information about future risk-adjusted returns is examined. In Section 6, we compare the atpRating with the Morningstar Rating and illustrate the differences in persistence of returns and costs. In Section 7, we analyze how a sub-sample of investors who have been exposed to the atpRating as well as the Morningstar Rating have chosen to invest in different mutual funds. Section 8 offers conclusions.

2. The Danish market for mutual funds

In this section, we provide a brief introduction to the Danish market for mutual funds after which we focus on the description of the costs associated with investing in Danish mutual funds. For the interested reader, Bechmann and Rangvid (2007) provide a comprehensive and detailed description of the Danish mutual fund market.

In our investigation, both equity funds and bond funds are looked at. Equity funds and the bond funds are divided into four categories following the classification used by the ATP (listed in Appendix A). The equity funds are divided into the following categories: *Danish Stocks*, *Global Stocks*, *Regional Stocks*, and *Other Stocks*. *Global Stocks* refers to mutual funds investing in several countries, i.e. global portfolios, whereas *Regional Stocks* refers to mutual funds holding stocks from individual countries or regions, for instance, US large cap funds, Europe large cap stocks, etc. The bond funds are divided into *Short Bonds*, *Long Bonds*, *Global Bonds*, and *Other Bonds*. *Short Bonds* and *Long Bonds* refer to Danish and Euro-zone bonds. These eight ATP

[5] Dahlquist et al. (2000) and Korkeamaki and Smythe (2004) are two of the few studies that evaluate the relation between costs and performance of non-US funds. Dahlquist et al. (2000) study Swedish mutual funds and Korkeamaki and Smythe (2004) Finnish funds. Both papers report that higher costs affect performance negatively. Christensen (2003, 2005), who also studies Danish mutual funds, focuses on issues other than the relation between mutual fund costs and performance. Additional studies of non-US mutual fund performances are Cai et al. (1997), Blake and Timmermann (1998), Otten and Bams (2002), Cesari and Panetta (2002), and Deaves (2004).

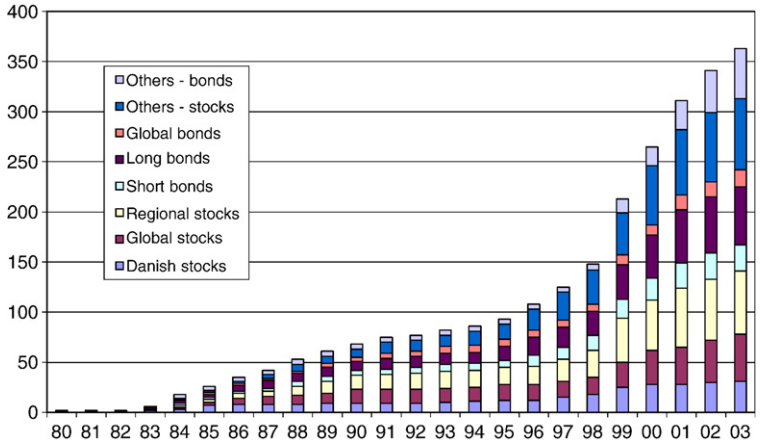

Fig. 1. Number of Danish mutual funds in different categories during the sample period.

categories combine different Morningstar categories (for instance the ATP category *Global Stocks* consists of the Morningstar categories Global Large Cap stocks and Global Mid/Small Cap stocks; see Appendix A). ATP uses fewer categories than Morningstar in order to reduce the dimensions of the investment universe faced by savers in the Special Pension Savings Scheme.

In Fig. 1, the development of the number of Danish mutual funds during the period from 1980 to 2003 is shown in order to provide a picture of the growth and composition of the Danish mutual funds market. In the figure, the funds are sorted into the eight ATP investment categories. Since 1995, the Danish market for mutual funds has gained momentum. For instance, in 1995 there were fewer than 100 Danish mutual funds, but less than ten years later there were more than 350 funds. Since 1990, the average annual growth rate of the number of Danish mutual funds has been approximately 14%. The total value of Danish mutual funds' holdings increased from DKK 21 billion in 1990 to DKK 257 billion in 2000, reaching a value of DKK 364 billion in 2003.[6] This corresponds to an average annual growth rate of approximately 24.5%. Given the fact that the total value of assets increased by 24.5% on an annual basis during the 1990–2003 period whereas the total number of funds increased by 14% on average, the average fund size has increased. Finally, to put the size of the Danish mutual fund market in perspective, notice that the total value of the Danish stock market in 2003 was DKK 718 billion, i.e. holdings (of Danish and foreign assets) by Danish mutual funds corresponded to approximately 50% of the value of the Danish stock market. In 2003, more than 600,000 Danes, and correspondingly around 25% of Danish households, had invested directly in Danish mutual funds.

There are more equity funds than bonds funds in Denmark, and this has been the case every year since 1985 (each year, approximately 60% of the funds have been equity funds). In 2003, there were 212 equity funds and 151 bond funds, i.e. all in all 363 funds. We start our investigation in 1994. In 1994, we have data on 86 funds. The size of our sample consequently ranges between 86 funds and 363 funds. In comparison to the literature studying mutual fund

[6] One Danish krone (DKK) corresponds to approximately USD 0.18, i.e. the total value of assets under management by Danish mutual funds was approximately USD 64.1 billion in 2003.

markets other than the US market (Dahlquist et al., 2000; Otten and Bams, 2002; Cesari and Panetta, 2002; Deaves, 2004; Christensen, 2003, 2005), a sample of between 86 and 363 funds is a relatively large sample.

3. Expenses

In this section, the costs investors have to pay when investing in Danish mutual funds and the cost data used in the analysis are described.

There are three main sources of costs in mutual fund investing: front-end load fees, back-end load fees, and operating expenses. Because we have data on load fees and operating expenses from 1994 and onwards, the period from January 1994 to December 2003 is investigated. Each cost source is described in turn below.

Front-end load fee: When buying a newly issued mutual fund share, a front-end load fee is normally charged. The load fee is denoted in percent and is a one-time, transparent up-front fee. The front-end load fee is what the investor is charged on top of the Net Asset Value (NAV) per share in order to cover the costs the fund has when it purchases new assets as a result of the arrival of new investors in the fund. The front-end load fee ensures that the existing investors in a fund are left unaffected when the fund adjusts its asset holdings due to the new investors in the fund. Furthermore, the front-end load fee includes remuneration from the mutual fund to the bank (or other financial intermediary) that has established the sale of the mutual fund share to a private investor.[7]

An important difference between the Danish market and, for instance, the US market, is that in Denmark all funds may charge front-end load fees (as well as back-end load fees, described below), i.e. no distinction is made between non-load and load funds in Denmark. Table 1 provides summary statistics on the size of expenses in Danish mutual funds in 2003. The average front-end load fees vary between 1.7% and 2.4% of the NAV for equity funds, whereas they are between 0.8% and 1.9% for bond funds, i.e. entry costs are higher for equity funds than for bond funds.[8]

Back-end load fee: When selling a mutual fund, a back-end load fee can be charged. The back-end load fee is the reduction in the Net Asset Value per share that the investor must accept when selling a mutual fund share. The back-end load fee covers the costs associated with the fund selling assets in order to pay back the investor. The fee is denoted in percent and is a one-time fee.

[7] In Denmark, mutual fund shares can also be traded on the Copenhagen Stock Exchange. Stocks in mutual funds traded on the stock exchange are not subject to a front-end (or back-end) load fee. Given the large inflows of money to the mutual funds mentioned in Section 2, funds have mostly issued new shares, implying that most investors have had to pay the front-end load fees during the sample period we investigate. We thus include front-end load fees in our analysis, well aware that some investors do not have to pay the load fee if they are able to find the relevant mutual fund share on the stock exchange. Another way of interpreting our approach is that we investigate the costs to the investor if she trades directly with the mutual fund, i.e. buys a newly issued mutual fund share from the mutual fund, and sells it to the fund when terminating the investment. Finally, it should be noted that if the investor trades on the stock exchange, she will have to pay bid–ask spreads, i.e. there are other kinds of costs involved when trading on the stock exchange than the costs considered here.

[8] In the US, there has been a change in the way mutual funds charge their expenses. For instance, Barber et al. (2005) report that the assets under management by equity funds that charge front-end load fees have declined considerably (from 91% of the total equity controlled by equity funds in 1962 to 35% in 1999). At the same time mean front-end load fees have dropped from more than 8% in 1962 to approximately 5% in 1999 (see Fig. 1 in Barber et al., 2005). In Denmark, front-end load fees have remained approximately constant throughout the sample period, i.e. in Denmark there has not been a tendency for front-end load fees to be reduced.

Table 1
Costs in different categories of Danish mutual funds, 2003

		Stocks				Bonds			
		Danish	Global	Regional	Others	Short	Long	Global	Others
Ope. costs	Average	1.19	1.36	1.13	1.47	0.67	0.60	0.85	1.00
	Median	1.03	1.32	1.24	1.42	0.55	0.55	0.72	0.99
	Minimum	0.48	0.27	0.39	0.73	0.11	0.15	0.22	0.12
	Maximum	2.09	3.39	3.35	2.64	2.00	1.10	3.08	2.15
Front-end fee	Average	1.97	2.14	1.70	2.38	0.82	1.03	1.37	1.85
	Median	2.10	2.35	2.03	2.37	0.89	1.22	1.45	1.98
	Minimum	0.25	0.50	0.25	0.40	0.00	0.00	0.30	0.20
	Maximum	2.61	3.00	3.00	3.50	1.50	1.75	1.75	2.91
Back-end fee	Average	0.70	0.78	0.75	0.96	0.22	0.21	0.34	0.52
	Median	0.75	0.85	0.75	0.90	0.14	0.15	0.30	0.50
	Minimum	0.20	0.20	0.25	0.38	0.00	0.03	0.15	0.00
	Maximum	1.36	1.50	1.60	1.75	0.75	0.92	1.00	1.75

This table shows summary statistics of the operating expenses for 2003, the front-end load fees for 2003, and back-end load fees for 2003. We have divided the individual mutual funds into four bond categories and four equity categories. For each category of mutual funds, we provide the average cost, the median cost, and the minimum and maximum costs that a mutual fund within the category has charged.

Unlike the front-end load fee, the back-end load fee is unknown when the investor buys a mutual fund share, as the fee can change from the day of purchase to the day of sale. The averages of the back-end load fees charged by Danish mutual funds have been fairly constant throughout the sample period (not shown). Table 1 shows that back-end load fees are generally higher for equity funds than for bond funds, as is also the case with front-end load fees.

Operating expenses: Operating expenses cover mutual fund employee salaries, rental of mutual fund offices, marketing expenses, and so forth. Operating expenses are expressed as an expense ratio, i.e. as a percentage of the wealth (NAV) of the mutual fund.[9] Table 1 shows that operating expenses are also higher for equity funds than for bond funds; for all categories of equity funds, average operating expenses make up more than 1% of NAV, whereas they make up less than 1% of NAV for all categories of bond funds. Fig. 2 shows the development over time of the average operating expenses for the different categories of funds analyzed. For all equity categories, average operating expenses have increased from 1997/1998 to 2003. For instance, the average operating expense ratio for mutual funds investing in *Danish Stocks* was 1.19% in 2003, up from 0.71% in 1998; an increase of approximately 67%. The same kind of pattern is also witnessed in the other equity categories. We investigated whether the increase in operating expenses is due to newer funds being more expensive than older ones. This is not the case, i.e. the average costs of funds have indeed increased.

For the bond categories, the pattern is slightly different. The average operating expenses for the *Global Bonds* funds and the *Other Bonds* funds follow the same pattern of increases as the equity

[9] Danish mutual funds publish information about their expense ratio. Other European funds normally publish the management fees. The exact definition of these costs varies from country to country. However, the expense ratios normally differ slightly from management fees as the expense ratio includes more costs than do management fees. For instance, the expense ratio includes distribution fees while management fees generally do not. Similarly, the expense ratio can be said to be backward looking (as it measures the expenses that have in fact been paid throughout a year) whereas the management fee is forward looking.

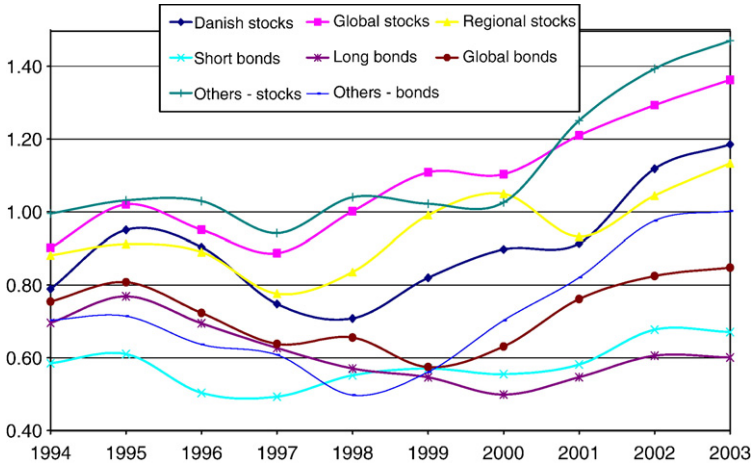

Fig. 2. Average operating expenses for Danish mutual funds, 1994–2003.

funds. The operating expenses of the *Short* and the *Long Bonds* funds, though, do not seem to have increased significantly during the same period.

Overall, as a last remark concerning the increase in operating expenses in recent years, it should be noted that even though the costs of investing in a Danish mutual fund have increased in recent years, the costs are still lower than in many other European countries (see e.g. Bechmann and Rangvid, 2007).

It should finally be mentioned that, in addition to the three main types of costs described above, there are some additional costs that we do not consider. For example, the investor may have to pay brokerage costs and a custody fee. Brokerage fees, however, depend more on the individual investor and her bank than on the mutual fund and, for this reason, we disregard brokerage fees in the analysis. Similarly, we do not examine custody fees since they normally only constitute a minor part of the total costs. Finally, we recognize that the turnover of a fund's asset has implications for the return to the investor (Chalmers et al., 1999). Unfortunately, we only have turnover data for the years 2000, 2001, 2002, and 2003. Furthermore, it is not easy to obtain the associated costs of turnover. Consequently, turnover is not considered in the present study.

4. Constructing the atpRating

In order to construct a cost rating, the different sources of costs involved in mutual fund investments must be weighted into one cost indicator. We will then use the size of this cost indicator for each individual fund to sort the funds within the different rating categories and finally to assign them an atpRating.

4.1. The cost indicator

As mentioned, there are three main sources of costs associated with mutual fund investments: front-end load fee, back-end load fee, and operating expenses. Each of these costs must be given

an appropriate weight so that an indicator can be constructed. The theoretical cost indicator CI of a fund in some year t thus takes the form

$$\mathrm{CI}_t = \gamma_1 \mathrm{Ope.}_{t-1} + \gamma_2 \mathrm{Front}_t + \gamma_3 \mathrm{Back}_t,$$

where $\mathrm{Ope.}_{t-1}$ is the operating expenses of the fund during year $t-1$, Front_t is the front-end load fee, and Back_t is the back-end load fee. The load fees are both measured as the values valid in year t. The γs are the weights to be put on the different components of the cost indicator.

Given that both one-time loads and continuous operating expenses are paid when investing in Danish mutual funds, the investment horizon will matter for the importance of the different costs involved in a mutual fund investment. For instance, the return for an investor with a short horizon is greatly affected by the load fees, whereas this is less true for an investor with a longer horizon. These facts imply that there is no uniformly correct way to determine the weights of the cost indicator, as the "correct" weights will change depending on the investment characteristics of the individual investor.

Thus, in the following, we first choose weights and later investigate how robust the atpRating is towards changes in the chosen weights. In our choice of weights, we recognize that:

(i) The back-end load fees that the investors may be charged when leaving the fund are not known at the time of the purchase of the mutual fund share (as mentioned in Section 3). On the other hand, front-end load fees are known when purchasing the fund.
(ii) Back-end load fees and front-end load fees are correlated, i.e. a large part of the information contained in back-end load fees is also contained in front-end load fees (the correlation across funds between the average back-end and front-end load fees over 2001, 2002, and 2003 is 0.52).

For these two reasons, we decided to put zero weight on the back-end load fee, and include the front-end load fee with a "double-weight" in the cost rating. Furthermore, we note that:

(iii) Operating expenses are paid each year, whereas load fees are paid only once.

For this reason, a higher weight is put on operating expenses. On the basis of these considerations, the baseline specification of the cost indicator we work with takes the form:

$$\mathrm{CI} = 0.7\mathrm{Ope.} + 0.3\mathrm{Front} + 0\mathrm{Back}.$$

In other words, operating expenses weight 0.7 in the costs indicator, front-end load fees weight 0.3, and the coefficient to the back-end load fee is set at zero. Calculations not reported show that the 30% and 70% weights generally correspond to an implicit investment horizon of approximately five years.[10]

In addition to our baseline specification, we also discuss results using a cost indicator where operating expenses weight 0.9 and front-end load fees weight 0.1 (approximately corresponding to an investment horizon of 15 to 20 years), as well as an indicator where operating expenses are

[10] In unreported calculations, we found that the present value of load fees for many of the categories we use make up approximately 30 percent of total costs with a five-year investment horizon.

Table 2
Numbers of funds in different ATP categories

	Stocks				Bonds				Total
	Danish	Global	Regional	Others	Short	Long	Global	Other	
3 years	28	35	50	60	21	43	12	19	268
5 years	21	17	28	34	15	26	–	–	141
8 years	12	16	17	17	–	14	–	–	76
10 years	–	14	17	–	–	11	–	–	42

This table shows the number of Danish mutual funds in each ATP category when the period to be analyzed spans 3 years (the 2001–2003 period), 5 years (the 1999–2003 period), 8 years (the 1996–2003 period), and 10 years (the 1994–2003 period). The table shows the number of funds in categories with at least ten funds. If there are less than ten funds in a category, the funds in the category are not rated in the atpRating. In the table, a dash indicates that there are less than ten funds in the particular category.

given full weight. Given the results of the investment horizons implicitly determined by the different weights, we have decided not to report any results for a cost indicator where the weight on operating expenses is lower than 70%. However, such cost indicators lead to results similar to those discussed later in this paper.

4.2. The atpRating

As the final step in the rating process, all funds are sorted according to the size of their cost indicator in relation to the cost indicators of the other funds in the same ATP category. One crown, the high-cost designation, is assigned to the 10% of the funds with the highest costs (as measured by the size of the funds' cost indicators) within their ATP category. The next 22.5% of funds within an ATP category are assigned two crowns. The next 35% with three crowns are the funds that have costs around the value of the average cost indicator of their ATP category. The next 22.5% are assigned four crowns, and the 10% of the funds with the lowest costs within an ATP category are assigned five crowns. The percentages used for the rating of the funds are the same percentages as Morningstar uses for the rating of funds when assigning stars to mutual funds.

Our total sample period is January 1994–December 2003, as mentioned. In our baseline examination, we rated the funds in 1994, 1995, 1998, and 2000, i.e. ten years before the last year in the sample, eight years before the last year in the sample, five years before the last year, and three years before the last year in the sample.[11] We decided that each ATP category should contain at least ten funds before it made sense to rank them. From Table 2, which shows the numbers of funds in the different ATP categories for the different time horizons, it can be seen that there are ratings within all categories using three years of data, whereas *Global Stocks*, *Regional Stocks*, and *Long Bonds* are the only groups that have ratings for the ten-year horizon.

4.2.1. Robustness of the atpRating towards changes in the weights

The first analysis of the atpRating that we perform is to examine how sensitive the ratings of the funds are with respect to the choice of weights in the cost indicator. Table 3 reveals how the ratings of funds in 2003 change when the weight put on operating expenses is altered from 0.7 to

[11] When analyzing ten-year future performance, we should in principle calculate the cost indicator and the ratings in 1993. Due to cost data availability, we ranked the funds in 1994, however.

Table 3
Sensitivity of fund ratings to the choice of weights in the cost indicator, 2003

		Weight on Ope. expenses = 0.9				
		1	2	3	4	5
Weight = 0.7	1	25	9	2	0	0
	2	10	39	25	2	0
	3	1	26	76	16	0
	4	0	2	15	46	11
	5	0	0	2	6	30

This table illustrates the effect on the atpRating of changing the weight on operating expenses from 0.7 to 0.9. For instance, the first row shows that 25 funds were ranked into atpRating category 1 regardless of whether the weight on operating expenses was 0.7 or 0.9. Nine funds that were initially rated in category 1 when the weight was 0.7, changed to category 2 when the weight was changed to 0.9, and 2 funds changed from category 1 to category 3 as a consequence of the change in the weight on operating expenses from 0.7 to 0.9.

0.9 (which corresponds to a change in the weight on operating expenses of 28.5%, i.e. a considerable change) and consequently the weight on the front-end load fee from 0.3 to 0.1.[12]

Table 3 reveals, for instance, that there are 25 funds that get an atpRating of one crown regardless of whether the weight on operating expenses is 0.7 or 0.9. Nine funds change from rating 1 to rating 2 when the weight is changed from 0.7 to 0.9, and two funds change from rating 1 to 3 when the weight is changed from 0.7 to 0.9. No fund changes from rating 1 to rating 4 or 5.

Altogether we find that 63% of the funds remain in the same rating category regardless of whether the weight on operating expenses is set at 0.7 or 0.9. Thirty-four percent of the funds change one rating category (either up or down), when the weight is changed from 0.7 to 0.9, and 3% of the funds change two rating categories. No fund changes more than two rating categories. Overall, 97% of the funds remain in their rating category or change one rating category.

If the weight of the cost indicator is changed from 0.7 to 1.0 (results not shown in order to save space), 93% of the funds remain in the same rating category or change one rating category (51% remain in the same category and 42% change one category).

One reason why so relatively few funds change rating category when the weights are changed is that the correlation between operating expenses and front-end load fees across funds is close to 0.5 for the 2001–2003 period. In other words, a fund that has a relatively high (low) front-end load fee is likely to have relatively high (low) operating expenses, too. Changing the weight to the front-end load fee, thus, will not change the relative sizes of the cost indicators of the funds much, i.e. the atpRating is rather robust with respect to the chosen weights in the cost indicator.

5. Does the atpRating predict future performance of the funds?

The atpRating rates funds according to the size of their costs and in this way summarizes information about the different kinds of costs investors have to pay in a transparent fashion. In addition to using the rating for this purpose, one would also like to know whether the information about costs that the rating contains is relevant for predicting future performance. The idea

[12] Due to the lack of data on operating expenses for 20 newly started funds, we are only able to rate 343 funds (out of a total of 363 funds) in 2003.

underlying this hypothesis is that if two funds generate similar pre-cost returns, the return to the investor, per definition, will be higher if the investor invests in a fund with relatively low costs today, and these costs are persistent such that low costs in the current period are associated with low costs in future periods (we evaluate the persistence of costs in Section 6.2).

5.1. Measuring fund performance

Before investigating the predictive content of the atpRating, we briefly explain how risk-adjusted returns are measured.

Given risk-return trade-offs, risk averse investors are not only interested in the return they receive, but also in the certainty with which they can expect to receive those returns. For this reason, it is now standard to look at risk-adjusted performance measures (a fund showing a high return could also be a risky fund — a possibility that is taken into account when using risk-adjusted performance measures). The standard performance measure used, and the one applied here, is Jensen's alpha, which is the constant α_i from the time series regression

$$r_i - r_f = \alpha_i + \beta_i(r_p - r_f),$$

where r_i is the total return to the investor from holding fund i, r_p is the return on a benchmark portfolio, r_f is the return on the risk-free asset (the one-month Copenhagen Interbank Offered Rate, CIBOR1m, as our return data are sampled monthly), and β_i is the beta of the fund with respect to the benchmark portfolio.

Mutual fund returns are defined as the percentage changes in the Total Net Asset Values of the funds plus dividends if they have been paid out. The Total Net Asset Value of a fund is the total value of the assets that the fund holds taking into account the operating expenses of the fund, i.e. *all else equal* higher operating expenses of a fund lead to lower Net Asset Values and thus returns.

A decision must be made regarding the benchmark portfolio and periods to be used when estimating the alphas and the betas. In the US literature, it is now customary to generate alphas from a four-factor model (Gruber, 1996; Carhart, 1997; Wermers, 2000), i.e. to use several different benchmark portfolios. In the US literature, however, focus is often restricted to US equity funds only. In this study, we have both equity funds, bonds funds, money market funds, and so forth. Furthermore, we have funds investing in Denmark only and funds that invest in other countries. In other words, applying the same model for determining expected returns of all funds would not be appropriate as the benchmark portfolios will vary from one mutual fund class to another.

The approach we follow below is to keep the model as simple as possible and regress each fund's excess returns on the excess return of one single benchmark portfolio. However, we also rely on the insight of the US literature that one cannot expect the same single benchmark portfolio to capture the many facets of returns from many different funds with many varying characteristics. For this reason, we regress the return from fund i on the return from the Morningstar category associated with fund i

$$r_i - r_f = \alpha_i + \beta_i(r_{p,j} - r_f),$$

where $r_{p,j}$ is the return from the Morningstar-defined category j portfolio to which fund i belongs. The major advantage of this approach is that the performance of fund i is measured in relation to the performance of its closely related peers, whereby we, here, follow the approach of

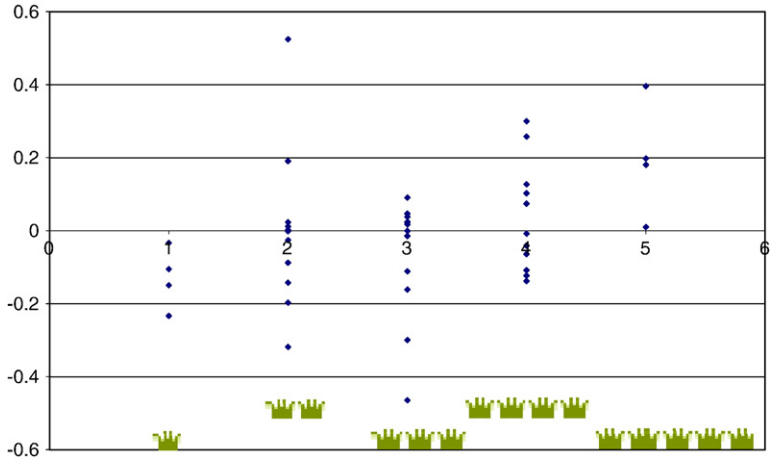

Fig. 3. Alphas of the individual funds estimated in the 1994–2003 period against the atpRating category of the individual funds in 1994. The alphas measure the average monthly risk-adjusted excess return obtained for the 1994–2003 period.

Morningstar, which also gives a fund stars based on the performance of the fund in relation to the performance of other funds within the Morningstar category; see Section 6.

It should be mentioned that the Morningstar categories constitute relevant benchmarks for the funds. For instance, the average R^2 across all regressions used to generate the alphas is 0.85 when estimating the models using three years of data, i.e. a Morningstar category return captures on average around 85% of the variation in the return of a fund during the 2001–2003 period. For the alphas estimated using five years of data, the average R^2 is 0.81 and with eight years of data the average R^2 is 0.77. All in all, we use 53 Morningstar categories as benchmarks. The categories and their summary statistics are listed in Appendix A.

We examine alphas based on estimations using three, five, eight, and ten-year periods of observations. Given that there is information in all alphas, whether they are statistically different from zero or not, we follow the standard approach taken in the literature (Gruber, 1996; Carhart, 1997; Dahlquist et al., 2000; Wermers, 2000 etc.) and use all available alphas in our cross-sectional investigations.

5.2. Testing the atpRating

Two kinds of analyses are presented in order to shed light on the question of whether the atpRating contains information about the future performances of mutual funds. First, the future performance of a fund is graphed together with the fund's atpRating. Second, the cross-fund regression approach of Blake and Morey (2000) is used to test whether the atpRating contains information about the future performance of the funds.

5.2.1. Graphically illustrating the predictive power of the atpRating

To illustrate the predictive power of the atpRating, consider 1994 as the base year. We rank and assign crowns to the funds based on the sizes of their cost indicators in 1994, as described in Section 4. We then plot the rating of a fund in 1994 against the alpha of the fund where the alpha is generated for the January 1994–December 2003 period. We are interested

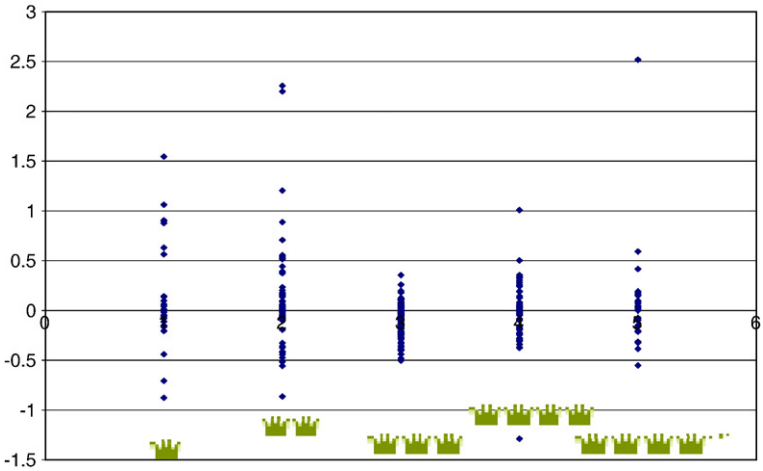

Fig. 4. Alphas of the individual funds, estimated in the 2001–2003 period, against the atpRating category of the funds in 2000. The alphas measure the average monthly risk-adjusted excess return obtained over the 2001–2003 period.

in evaluating whether the alphas, based upon the 1994–2003 period, on average are higher for those funds that in 1994 were assigned, for instance, five crowns (i.e. the 10% of the funds with the lowest costs in 1994), compared to the average alpha of those funds that were ranked, for instance, with one crown in 1994. We perform the same exercises using 2000 as starting year.

In Fig. 3, we present the alphas of the funds when measured over the 1994–2003 period against their atpRating in 1994; the alphas are monthly alphas, i.e. they measure the average monthly risk-adjusted excess return obtained for the 1994–2003 period. Fig. 3 reveals a clear tendency that funds with higher ratings in 1994 (and thus lower costs in 1994) have done better during the 1993–2004 period than funds with lower ratings (and thus higher costs), perhaps with the exception of the large dispersion of alphas for the funds in atpRating category 2. In particular, funds assigned three crowns in 1994 seem to have higher alphas than funds assigned one crown in 1994; funds in group four, on average, did better than those in group three, and funds in group five, on average, did better than those in group 4. In order words, there seem to be indications — on the basis of this first graph — that funds with higher ratings (i.e. lower costs) yield a better future long-run (ten-year) performance than do funds with lower ratings. Indeed, the average alpha of funds in group five (the funds with the lowest costs in 1994) was 0.20 whereas the average alpha of funds in group one was −0.13, a difference in average alphas that is statistically significant at the 1% level. In other words, the low-cost funds have outperformed the high-cost funds by around 0.33% per month.

Fig. 4 is constructed like Fig. 3. However, in Fig. 4, three-year alphas are shown together with the atpRating of the funds in 2000. The figure reveals that there is no clear pattern in the alphas when comparing alphas of lower rated funds with those of higher rated funds. In other words, there seems to be less information about future returns over short horizons than there is when evaluating a fund's performance over long horizons (as in Fig. 3); indeed, the differences in the respective average alphas of the high-cost and low-cost funds are not statistically significant. There are funds in rating category 1 (i.e. those with the highest costs in 2000) that have performed very well, but there are also funds in category 1 that have obtained very low alphas. In other

Table 4
Dummy variable regression. All funds

	Obs	Adj. R^2	p-value	δ_1		δ_2		δ_3		δ_4		δ_5	
			F-test	Cons	p-value	Rat 2	p-value	Rat 3	p-value	Rat 3	p-value	Rat 5	p-value
3 year, OLS	268	0.077	**0%**	0.213	**0%**	−0.108	22%	−0.323	**0%**	−0.229	**1%**	−0.127	22%
3 year, Robust				−0.026	56%	0.053	33%	−0.074	15%	−0.007	90%	0.022	73%
5 year, OLS	141	0.051	13%	0.152	**6%**	−0.144	14%	−0.156	**9%**	−0.160	10%	0.032	78%
5 year, Robust				−0.010	86%	−0.033	63%	−0.021	74%	−0.043	53%	0.045	58%
8 year, OLS	76	0.163	**1%**	−0.008	92%	0.010	92%	−0.032	74%	−0.064	53%	0.295	**2%**
8 year, Robust				−0.103	14%	0.067	42%	0.046	56%	0.055	50%	0.248	**1%**
10 year, OLS	42	0.193	**9%**	−0.130	14%	0.110	29%	0.080	43%	0.165	11%	0.326	**1%**
10 year, Robust				−0.130	11%	0.064	50%	0.118	21%	0.153	11%	0.322	**1%**

This table shows the results of the regressions of alphas of Danish mutual funds on a dummy variable that picks out funds in atpRating category 2 (the estimated coefficient to the dummy being δ_2), a dummy that picks out funds in atpRating category 3 (the estimated coefficient being δ_3), for funds in category 4 (coefficient δ_4), and for funds in category 5 (with estimated coefficient δ_5). δ_1 is the estimate of the constant. The table shows the estimated coefficients and the associated p-values from t-tests of whether the particular coefficient is statistically distinguishable from zero. The coefficients have been estimated using simple OLS techniques and robust estimation techniques that take into account the influence of outliers. Finally, the table shows the number of observations in each regression and the R^2s.

words, high costs do not necessarily generate high performance. This is in accordance with the findings in Gruber (1996) for the US. But high costs do not necessarily imply low short-run performance either. Instead, high costs are not only associated with funds that have generated a high future risk-adjusted short-run (three-year) performance, but also with funds that have generated a very low future risk-adjusted short-run performance. We will discuss potential reasons for the differences in results that are obtained when using three years of data versus ten years of data in Section 6.2.

5.2.2. Testing the atpRating using regression analyses

An alternative way to more formally test the predictive power of the atpRating is to follow the approach of Blake and Morey (2000) and run multivariate cross-sectional regressions of alphas on dummy variables that pick out four of the five ATP categories. The advantage of this approach is that it uses all the funds in each single regression.

The regressions are undertaken using the ratings of the funds in 1994, 1995, 1998, and 2000, where the ratings in the individual years are obtained as explained in Section 4. For a given year, we create a dummy variable that picks out the funds with two crowns, another dummy variable that picks out the funds with three crowns, etc. No dummy variable is created that picks out the funds with one crown (the funds with the highest costs), i.e. the funds with one crown are used as

the reference funds with which other lower-cost funds are compared. To explain the procedure, consider the rating of funds in 1994. We regress the alphas of the funds on the four dummies, remembering that the alphas are based on the period from 1994 to 2003, i.e. on the risk-adjusted performance of the fund in the ten years following the rating. In this way, the regressions provide evidence of the information about future returns that the atpRating contains. The same procedure is followed using 1995, 1998 and 2000 as the starting year, where the alphas are estimated over the following eight, five and three years. The regression we perform for the full sample of funds looks as follows:

$$\alpha_i = \delta_1 + \delta_2 D_2 + \delta_3 D_3 + \delta_4 D_4 + \delta_5 D_5,$$

where D_2 is a dummy picking out those funds that belong to atpRating group 2 (the group of funds with the second-highest costs), D_3 picks out those funds that belong to rating group 3, and so forth, up to D_5 which picks out the funds with the lowest costs within their ATP category. If the atpRating contains information about future returns, we expect that funds with lower costs also have higher alphas (perform better), i.e. we expect δ_2 to δ_5 to be positive, and we expect δ_3 to be larger than δ_2, δ_4 to be larger than δ_3, and δ_5 to be larger than δ_4.

The results from these regressions are presented in Table 4. We present results using both simple OLS regressions and robust regressions. Results from robust regressions are presented in order to check the robustness of the OLS regressions with respect to the influence of outliers (see, for example, Holland and Welsch, 1977).

The results contained in Table 4 (again) reveal that rating funds on the basis of their cost indicator provides information about the risk-adjusted excess returns of the funds over longer investment horizons. In more detail, looking at the OLS regression results for risk-adjusted returns over the three years as based on the values of the cost indicator in 2000 (the estimates presented in rows "3 year"), all coefficients have the "wrong" sign in the sense that funds with lower costs than those of group one also experienced lower risk-adjusted returns, i.e. the opposite of what would be expected. The same basically holds true for the five-year returns. Furthermore, signs and significance levels change if estimation techniques are changed, indicating that some outliers influence the results, see also Fig. 2.

For the eight and ten-year returns, however, the story is different. In particular, the rating of funds in 1995 implied that funds in group 5 (the funds with the lowest costs in 1995) actually obtained statistically significant higher risk-adjusted excess returns for the 1996 to 2003 period than funds with the highest costs in 1995 (those in group 1). In other words, on average, the funds with the lowest costs performed better than the funds with the highest costs. The rating of funds in 1994 also had some predictive power. The funds in the group with the lowest costs, group 5, obtained significantly higher risk-adjusted excess returns over the out-of-sample period 1994–2003 than the funds in group 1 (with the highest costs). Furthermore, for the eight and ten-year regressions, the results are generally robust towards the influence of outliers, as the OLS and robust estimation techniques generally give similar results.

What is the interpretation of the coefficients? Consider the estimate of δ_5 as an example. The estimate of δ_5 gives the difference between the alphas generated by funds in groups 1 and 5 respectively, i.e. an estimate of δ_5 of 0.326 implies that the average monthly alpha of the funds in group 5 is 0.326% higher than the average alpha of the funds in group 1. In other words, had an investor in 1994 invested in the funds in group 5 (those with the lowest costs in 1994 within their ATP category), she would have obtained an annual risk-adjusted excess return that is

Table 5
Dummy variable regressions over different subsamples

	Obs	Adj. R^2	*p*-value	δ_1		δ_2		δ_3		δ_4		δ_5	
			F-test	Cons.	*p*-value	Rat 2	*p*-value	Rat 3	*p*-value	Rat 4	*p*-value	Rat 5	*p*-value
3 year regressions													
1994–1996	42	0.104	38%	−0.132	31%	0.135	37%	0.157	29%	0.169	26%	0.369	**5%**
1998–2000	119	0.011	87%	−0.064	34%	0.006	94%	0.053	49%	0.052	52%	0.042	66%
2001–2003	*268*	*0.077*	**0%**	*−0.026*	*56%*	*0.053*	*33%*	*−0.074*	*15%*	*−0.007*	*90%*	*0.022*	*73%*
2003–2005	327	0.014	34%	0.036	40%	−0.059	26%	−0.087	**8%**	−0.079	13%	−0.003	96%
5 year regressions													
1994–1998	42	0.081	53%	−0.158	19%	0.092	51%	0.167	22%	0.172	22%	0.265	12%
1998–2002	119	0.020	68%	−0.061	28%	0.076	27%	0.070	27%	0.018	79%	0.080	32%
1999–2003	*141*	*0.051*	*13%*	*−0.010*	*86%*	*−0.033*	*63%*	*−0.021*	*74%*	*−0.043*	*53%*	*0.045*	*58%*
2001–2005	254	0.047	**2%**	−0.033	44%	0.034	51%	−0.058	24%	−0.002	97%	0.006	92%
8 year regressions													
1994–2001	42	0.138	23%	−0.115	24%	0.079	48%	0.100	37%	0.135	24%	0.306	**3%**
1996–2003	*76*	*0.163*	**1%**	*−0.103*	*14%*	*0.067*	*42%*	*0.046*	*56%*	*0.055*	*50%*	*0.248*	**1%**
1998–2005	109	0.109	**2%**	−0.102	**4%**	0.081	18%	0.070	21%	0.014	82%	0.173	**1%**
10 year regressions													
1994–2003	*42*	*0.193*	**9%**	−0.130	11%	0.064	50%	0.118	21%	0.153	11%	0.322	**1%**
1996–2005	72	0.131	**5%**	−0.073	28%	0.015	86%	0.007	93%	−0.017	83%	0.138	15%
12 year regressions													
1994–2005	41	0.186	11%	−0.135	**8%**	0.082	37%	0.085	33%	0.138	13%	0.275	**1%**

See notes to Table 4. The results here are from the robust regressions as also presented in Table 4. The emphasized estimates are those from the original sample period 1993−2003, i.e. the estimates from Table 4.

approximately 3–4% higher than if she had invested in funds in group 1 (with the highest costs within their ATP category in 1994) measured over the period from 1994 to 2003. Our findings are thus economically significant, too.

We find that the coefficient to the dummy picking out the lowest cost funds (δ_5) is significantly positive using the 1994–2003 out-of-sample period. We also find that the coefficients δ_2, δ_3, and δ_4 are estimated to be positive with regard to long horizons, but that they are not significantly different from zero (δ_4 is significant at a 11% level, however). The finding that the lowest-cost funds have a significantly superior long-run out-of-sample performance in relation to the highest costs funds, but also that funds in groups 2, 3, and 4 do not have significant coefficients (even when the signs of the coefficients are positive and the δs are increasing in magnitude) is in line with the results of Blake and Morey (2000) on the predictive content of Morningstar Ratings. Blake and Morey (2000) report that Morningstar Rating groups 1 and 2, i.e. the rating groups picking out the funds with the historically lowest returns within their Morningstar categories, predict low risk-adjusted returns out-of-sample in comparison to the return of funds in Morningstar Rating group 5 (with the historically highest returns), whereas Morningstar Rating groups 3 and 4 have no predictive power.

5.2.2.1. Subsample analyses and an extension of the sample. We have shown that the atpRating predicts performance over longer horizons but not over shorter horizons. In order to get a closer look at whether this is a systematic finding over different subsamples, we present in Table 5 results from the same kind of regressions as those in Table 4, but for different subsamples. In addition, we extend our baseline sample period by two years so that the data cover the period until 2005. It should be mentioned, though, that the classification of the funds into their relevant Morningstar categories is not perfect for the 2004–2005 period. We will therefore view the results from the samples extending up to 2005 as approximate only, which is also why we focus on the January 1994–December 2003 sample elsewhere in the paper.[13]

Table 5 reveals that the finding that the atpRating does not predict future performances on the shorter horizons, but does so on the longer horizons, is quite robust across different subsamples. In particular, the coefficients from the regressions relating the ranking of a fund in period t to the fund's risk-adjusted returns over periods t to $t+3$ or $t+5$ are generally insignificant (there are some significant coefficients, but there is no systematic pattern in the significant coefficients). On the other hand, regardless of the period over which the eight year risk-adjusted returns are measured, the coefficient to δ_5 is significantly positive, i.e. if the investor had bought a low-cost fund in 1994, 1996, or 1998, the investor would on average have received a higher risk-adjusted return over the following eight years, compared to buying a high-cost fund in 1994, 1996, or

[13] To illustrate why the classification of the funds into their relevant Morningstar categories is not perfect for the 2004–2005 period, consider the Morningstar category England Large Cap stocks. In the first version of this paper, the data extended to December 2003, and funds investing in UK stocks were associated with the England Large Cap stocks. When updating the data until December 2005 for use in the present version of the paper, the Morningstar England Large Cap stocks category had been split into three categories: England Large Cap Value stocks, England Large Cap Blend stocks, and England Large Cap Growth stocks. To deal with this issue, we use the England Large Cap stocks as a category until December 2003 (which is the latest observation in our original data). Afterwards, UK equity funds are split into the three categories mentioned above. We thus view the results for the 2004–2005 sample as approximate only, because this break in the categories will have a slight influence on the alpha estimates.

1998. In addition, there are 109 funds in the 1998–2005 sample, i.e. these (approximate) results pertain to a fairly high number of funds. Finally, for the two ten-year periods and the twelve-year period, δ_5 is also positive and generally significant.[14]

5.2.2.2. Sharpe's ratios as performance measures. We use Jensen's alpha as our measure of performance. By using the return on the Morningstar category to which a fund belongs as the relevant benchmark against which the performance of the fund is measured, we relate the performance of a fund to its direct peers. Even though we find this to be a reasonable way of measuring performance, our approach, of course, does not free us from relying on an appropriate benchmark. In order to examine the sensitivity of our results towards the benchmarks we use, we also analyzed the predictive information contained in the atpRating using Sharpe's ratios as performance measures. Given that Sharpe's ratios are calculated by dividing the excess return of a fund with its standard deviation, Sharpe's ratios are obviously not dependent on any benchmark index.

Consider a comparison of the average Sharpe's ratios first. For the 42 funds that we have data for during the 1994–2003 period, the average (median) Sharpe for rating 1 (the highest cost funds) is 0.01 (0.01) and 0.10 (0.10) for rating 5 (the lowest cost funds) — the difference being significant at the 5% level. A similar but less pronounced pattern is found if the eight-year data are examined. The pattern of higher Sharpe's ratios for the lowest costs funds in comparison to the highest costs funds is also found in the approximate data during the 1996–2005 ten-year period, i.e. for the 72 rated funds existing over the 1996–2005 period, the average (median) Sharpe for rating 1 is 0.14 (0.12) and 0.21 (0.17) for rating 5 — the difference being significant.

We also undertook regressions such as those in Table 5. The results from these analyses are in many ways similar to the results presented in Table 5.[15] In particular, over the three- and five-year horizons, most coefficients are insignificant and there are no systematic patterns in the few significant results. Over the longer horizons, however, the estimates of δ_5 are generally significantly positive, i.e. our overall results are fairly robust towards the use of other measures of performance.

5.2.2.3. Other weights in the cost indicator. We conducted the same tests as those in Table 5 using the alternative cost indicators where operating expenses weight 0.9 and front-end load fees 0.1, CI = 0.9Ope. + 0.1Front, and another indicator that takes the form CI = 1Ope. + 0Front. Overall, the results from these regressions are qualitatively comparable to those in Table 5, which should not be surprising given the fact that Table 3 revealed that 97% of the funds change at most one rating category when the weight on operating expenses is changed from 0.7 to 0.9.

5.2.2.4. More funds on the long horizon. In this paper we are mainly interested in the performance of the atpRating. For this reason, in our original 1994–2003 sample, we limit ourselves to the 42 funds that can be rated in 1994 (see Section 4.2). However, all in all, there are

[14] We also conducted regressions such as those in Table 4 for those 42 funds that are available during the original 1994–2003 sample period, in order to examine whether these 42 funds are in some way special. The results showed that the estimates were generally all insignificant at the three-year horizon, whereas they were significant and positive for long horizons, as in Tables 4 and 5, i.e. the result that the atpRating predicts risk-adjusted performance over longer horizons but not over short horizons is also found when examining only the 42 funds that are available over the complete 1994–2003 period.

[15] In order to save space, tables documenting these results are not shown. The results can of course be obtained upon request.

68 funds for which we have data throughout the complete 1994–2003 period. Given the restriction mentioned in Section 4.2 that there should be at least ten funds in an ATP category before it makes sense to rate the funds within a category, there are thus 26 funds we do not use. If the aim is to answer the overall question of whether low-cost funds on average generate higher returns than high-costs funds, an alternative approach could be taken where all funds are used. An evaluation could, for instance, be made of whether the 20% of funds with the lowest costs in 1994 were able to generate higher risk-adjusted returns on average during the 1994–2003 period than the 20% of funds with the highest costs in 1994 were able to generate in the subsequent years. We find that the difference in average (median) alpha is 0.13 (0.20) and that these differences are significant at the 5%–10% level. If we compare at the 50% of funds with the lowest cost with the 50% of funds with high cost, we also find that this difference is significant. All in all, we conclude that when we include all funds, and are thus able to expand the sample to more than the 42 funds that are rated for the complete 1994–2003 period, we still find that, on average, it has been a better strategy out-of-sample for the investor to invest in the low-cost funds in 1994.

6. Comparing the atpRating with the Morningstar Rating

We have developed a new rating for mutual funds. It is consequently of interest to know how the atpRating distinguishes itself from other ratings that are available for Danish mutual funds, i.e. the Morningstar Rating. The Morningstar Rating is based on historical performance measured in terms of the risk-adjusted returns of mutual funds. Morningstar rates Danish mutual funds based upon the previous three years of risk-adjusted performance.[16] The risk-adjusted performance of a Danish mutual fund is compared to the risk-adjusted performance of other European funds within the same Morningstar category. If the fund belongs to the 10% of European funds within the category that has obtained the best performance, the fund receives five stars. If the fund belongs to the next best 22.5% of European funds, the fund receives four stars and so forth. The percentages used for classifying funds into the different categories in the atpRating are thus borrowed from the Morningstar percentages.

The important differences between the atpRating and the Morningstar Rating are thus as follows:

- Morningstar rates Danish mutual funds on the basis of the historical (the previous three years) risk-adjusted returns of a fund. The atpRating rates funds on the basis of the previous single year's fees.[17]
- Morningstar compares the performance of Danish mutual funds with the performances of other European funds within the same Morningstar investment category. The atpRating compares the costs of Danish mutual funds with those of other Danish mutual funds within the same ATP investment category.

[16] For ratings of US mutual funds, Morningstar Ratings are based upon the historical performance of a fund during three different periods: the last ten years, the last five years, and the last three years (see e.g. Morey, 2002). For each period, the fund receives its Morningstar stars. The overall Morningstar Rating of a fund is then based upon a weighted average of the ratings of the fund for the three periods. If a fund in the US has existed for less than five years, but more than three, the overall Morningstar Rating of the fund is the rating based upon the last three years. It is this last Morningstar Rating strategy that is used for all Danish mutual funds, even though some Danish mutual funds have existed for more than ten years.

[17] As mutual fund returns are obtained after operating expenses, Morningstar's risk-adjusted performance measures implicitly subsume operating expenses. In other words, in the atpRating, the size of fees is the only parameter upon which funds are rated. In Morningstar, fees are implicitly included in combination with risk-adjusted returns when rating funds.

Table 6
The rating of funds by the Morningstar Rating and the atpRating in 2003

		atpRating (2003)				
		1	2	3	4	5
MS(2003)	1	0.0%	6.0%	5.5%	0.0%	0.0%
	2	4.2%	20.0%	20.9%	10.7%	23.5%
	3	37.5%	40.0%	48.4%	50.0%	35.3%
	4	29.2%	22.0%	17.6%	33.9%	17.6%
	5	29.2%	12.0%	7.7%	5.4%	23.5%
		100.0%	100.0%	100.0%	100.0%	100.0%

This table shows the relation between the ratings of mutual funds in 2003 by the Morningstar Rating (MS) and the atpRating. In each column, the distribution of the funds in a certain atpRating category across the different Morningstar categories is shown.

Morningstar only started rating Danish mutual funds in 2001, i.e. data on Morningstar Ratings of Danish mutual funds exist from 2001. As a result, it is not possible to do any serious analyses of whether Morningstar predicts future performance given this short period of information on Morningstar Ratings. What we can do, however, is to evaluate whether the atpRating and the Morningstar Rating contain the same information, or whether the two ratings actually contain different information.

Table 6 shows the distribution of funds across the two ratings in 2003. Consider, for instance, funds that have been assigned atpRating 1 (the 10% of funds with the highest costs) in 2003. Table 6 shows that the Morningstar Rating did not assign any of these funds to its category 1 (the "worst performance" category). Indeed, 4.2% of the funds that were assigned category 1 in the atpRating were assigned category 2 in Morningstar, 37.5%, category 3 in Morningstar, 29.2%, category 4 in Morningstar, and the remaining 29.2% were assigned category 5. In other words, 29.2% of the funds that had the highest costs in 2003 were also in the Morningstar category of funds that have had the best risk-adjusted performance for the last three years. We have performed the same exercise for the relation between fund rankings in 2001 (not shown to save space). The overall picture shows again that the majority of the funds are not placed in the same category in the atpRating and the Morningstar Rating. The investor thus gets different information when considering the cost-based atpRating, based upon last year's costs, versus the performance-based Morningstar Rating, which rates funds based on the last three years of risk-adjusted performance.

The finding that funds with high costs (atpRating 1) can be placed in the Morningstar categories of funds with the very best performance, but also in categories of funds with rather poor performance (e.g. in Morningstar category 2), links up well with the findings presented in Fig. 4 on the relation between the atpRatings in 2000 and the future alphas calculated using data from the 2001–2003 period. From Fig. 4, we conclude that some funds with high costs have performed well on short horizons, while other funds have performed badly. The relation between Morningstar Ratings and atpRatings in this section mirrors this finding.[18]

[18] Another aspect worth paying attention to when reading Table 6 is that few funds are assigned Morningstar Rating group 1. The explanation is that Danish funds are traditionally rated fairly high in European comparisons. For instance, in December 2003, the average number of stars assigned to Danish funds was 3.16, which was the second-highest average among 14 European countries.

6.1. Which rating should the investor rely upon?

In the end, the answer to the question of which rating to rely upon must be based on the empirical evidence. In Tables 4 and 5, we have presented evidence that the atpRating contains information about future risk-adjusted returns over long horizons (eight to ten years), and, in particular, for these horizons, funds with the lowest costs have yielded higher performance than funds with higher costs. Given the short history of Morningstar Ratings in Denmark, it is not yet possible to present analyses of whether Morningstar Ratings contain information about the short- or long-horizon future performance of Danish mutual funds.

What we can do, however, is to refer to the literature on whether Morningstar Ratings contain relevant information about the future performance of US mutual funds, where longer periods of observations on fund performances are available. This issue has been examined in Blake and Morey (2000), Morey (2002), and Vinod and Morey (2002). These authors find that Morningstar is good at predicting losers, but less successful at predicting winners.

A fund receiving only a few Morningstar stars has a tendency to perform poorly in the future, whereas the highest rated funds often do not perform better than three or four star funds. As mentioned, the findings on the predictive performance of the atpRating are that the atpRating predicts superior performance of its highest-rated funds over investment horizons of eight to ten years. Therefore, it will be interesting in the future to compare the predictive power of Morningstar Ratings with those of the atpRating as well as to examine whether the use of both ratings can improve the identification of mutual funds that will perform well.

6.2. Why are there differences between the ratings?

The Morningstar Rating rates funds on the basis of the last three years of risk-adjusted returns where these risk-adjusted returns are calculated after implicitly taking into account costs, whereas the atpRating is based solely upon last year's costs.

One reason why there are differences between the two ratings is that costs are more persistent than returns and market-adjusted returns. In this section, we briefly examine, in particular, the variation over time of operating expenses and market-adjusted returns by relating across funds expenses lagged once with current expenses and market-adjusted returns lagged once with current market-adjusted returns. Market-adjusted returns are measured as the return in excess of the return of a fund's peers (measured by the return on the Morningstar category to which the fund belongs).

The variations are illustrated in Figs. 5 and 6. At first glance, the figures give a strong indication of the high persistence in expenses (expenses from the previous and current year are scattered closely around the regression line) and a very low persistence of market-adjusted returns (market-adjusted returns from the previous and current year are spread all over the regression line). Another way of illustrating this is by presenting the results from the regression of last year's expenses on this year's expenses across funds and years. This regression generates the following result (with t-statistics in parentheses below):

$$\underset{}{\text{Expense}_t} = \underset{(13.62)}{0.18} + \underset{(72.52)}{0.89}\ \text{Expense}_{t-1},$$

with an R^2 of 0.77. The corresponding regression for market-adjusted returns is

$$(r_t - r_{p,j,t}) = \underset{(1.13)}{0.002} + \underset{(1.11)}{0.026}(r_{t-1} - r_{p,j,t-1}),$$

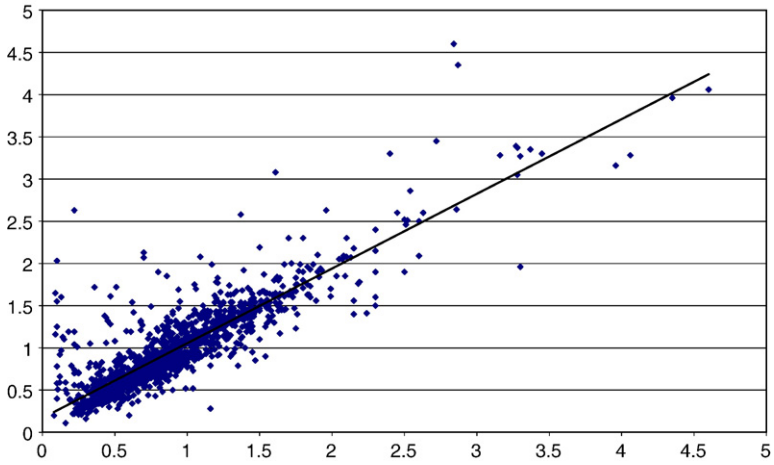

Fig. 5. Scatter plot of last year's operating expenses against this year's operating expenses.

with an R^2 of 0.00. These results thus indicate that the persistence in expenses is highly significant, and a large portion of the cross sectional variation in this year's expenses can be explained using last year's expenses (77%), whereas this is not the case for returns (we obtain similar results if marked-adjusted returns for, for example, a three-year period are considered). The finding that returns from Danish mutual funds are non-persistent is reported in Christensen (2005), too.

Given that returns are so volatile on an annual basis but costs are stable, it is difficult to imagine that one would be able to find significant relations between costs and returns in the short run, or, for that matter, between returns in subsequent years. On an annual basis, returns in year $t+1$ will fluctuate a great deal in relation to returns in year t. And given that costs do not fluctuate much, it cannot be expected that there is much of a relation between costs and returns in the short run either,

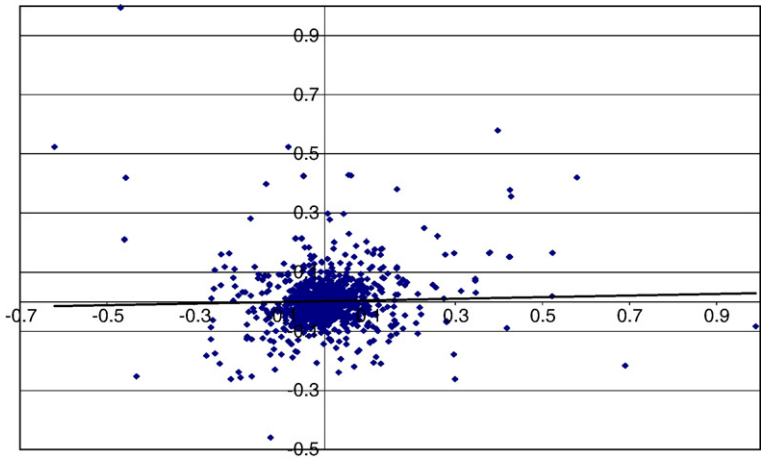

Fig. 6. Scatter plot of last year's return against this year's return.

as is also verified in Section 5. On the other hand, costs are very persistent, i.e. a fund that has high costs one year will most likely also have high costs in the following years. When accumulating these costs over many years, they will have a negative impact on returns, i.e. the influence of costs on performance will probably only be clear after several years — which is indeed what our results so far have indicated.

7. Do investors use the information contained by the ratings?

As mentioned in the introduction to this paper, the atpRating was originally constructed in order for it to be used when Danes, as of January 1, 2005, were given access to freely allocate their savings in the Special Pension Savings Scheme to different mutual funds. In this section, we describe the asset allocation of individuals who have chosen to actively allocate their savings between the different available mutual funds during the January 1, 2005–June 30, 2005 period. In doing so, we also contribute to the literature that analyzes a related case: the introduction of the Swedish system in 2000 allowing Swedes to allocate part of their pension savings (2.5% of their annual earnings) to different mutual funds. The Swedish system is analyzed in Cronqvist and Thaler (2004), Cronqvist (2006), and Karlsson et al. (2006).

Danes paid during the 1998–2003 period 1% of their salary to the Special Pension Savings Scheme, i.e. Danes have only been able to allocate a very small fraction of their overall pension savings freely as of January 1, 2005.[19] In addition, if a pension saver in the Special Pension Savings Scheme decides not to allocate her savings actively, these savings are invested in a default fund, as in the Swedish system described in Cronqvist and Thaler (2004), Cronqvist (2006), and Karlsson et al. (2006). The low value of savings that can be freely invested, and the fact that if they are not invested actively, they are allocated into a default fund, most likely also explains why so relatively few have chosen to actively allocate their funds into different mutual funds instead of just leaving their savings in the default fund (on June 30, 2005, i.e. six months after the opening up of the system, 4252 out of the total of 2.7 million Danes that have the possibility of allocating funds in the Special Pension Savings Scheme had invested actively). Concerning the amounts invested, as of June 30, 2005, approximately DKK 77 million, or roughly USD 14 million, have been actively allocated to the different mutual funds available. On average, each individual that has made active investment decisions has thus allocated approximately DKK 18,000.

Even though the amounts invested in the Danish Special Pension Savings Scheme are small, there is nevertheless one particular reason why it is interesting to study these active investment decisions: When allocating pension savings in the Special Pension Savings Scheme to different mutual funds, the investor has to use an Internet-based trading platform. The interesting feature of this trading platform is that it shows the Morningstar Ratings for the funds as well as the atpRatings. In other words, we know that when investors decide to allocate their funds into different mutual funds, they have seen both the Morningstar Rating and the atpRating for the funds. To be precise, the Internet trading platform shows the list of available mutual funds together with certain characteristics for each fund, such as the year-to-date return of the fund (*år til dato*), a description of the riskiness of the investment strategy of the fund (*risiko*), and so forth. We show a snapshot of the Internet trading platform

[19] The payments to the Special Pension Savings Scheme have been suspended during the 2004–2007 period. It is unclear whether Danes will have to pay to the Special Pension Savings Scheme again after 2007.

Fig. 7. The trading platform for investing the Special Pension Savings Scheme.

in Fig. 7. As can be seen from Fig. 7, the last two columns show first, the atpRating of a fund, and second, the number of stars a fund has been given by Morningstar (in the column "Rating"). Given that we know that investors have been exposed to both ratings (Morningstar and the atpRating), it is of interest to study how investor decisions are related to the ratings of the funds.

Before doing so, it should be mentioned that the investor, at any point in time, can click on the "?" icon in the columns "atpRating" and "Rating" in Fig. 7 to get a short description of the atpRating (that it rates funds based on past costs) and the Morningstar Rating (that it rates funds based on past performance). It should also be mentioned that some advertisements were conducted in Danish newspapers and television prior to January 1, 2005, by both the ATP and the funds themselves. These advertisement campaigns, though, were considerably smaller than the campaigns in Sweden associated with the opening up of the Swedish system (Cronqvist, 2006). Nor did Danes receive a physical book with the list of available funds, as Swedes did before the opening up of the Swedish system. On the other hand, the atpRating was heavily debated in the Danish press surrounding January 1, 2005, most likely because some funds were not happy about seeing the sizes of their costs being exposed so clearly. By following this debate in the press, investors also got information on the atpRating and its perceived strengths and weaknesses. Finally, the atpRating made its way into the Danish business press shortly after January 1, 2005.

Table 7 shows the first results regarding how investors have allocated their pension savings in relation to an allocation based on the funds' empirical distribution between the

Table 7
Allocations to mutual funds in different atpRating and Morningstar Rating categories. The first six months of 2005

	1	2	3	4	5	1+2	4+5
atpRating							
Empirical distribution	8.67%	21.97%	35.84%	22.54%	10.98%		
Over-/underweight (savings)	3.09%	−9.97%	−13.21%	15.20%	4.89%	−6.87%	20.08%
Over-/underweight (investors)	0.79%	−8.51%	−12.76%	12.08%	8.40%	−7.72%	20.48%
Morningstar							
Empirical distribution	3.42%	19.18%	30.14%	26.03%	21.23%		
Over-/underweight (savings)	2.42%	−12.70%	−2.19%	−2.92%	15.39%	−10.27%	12.46%
Over-/underweight (investors)	0.88%	−12.82%	−4.53%	−1.95%	18.41%	−11.94%	16.47%

This table shows the relative fractions of total savings (in rows "Over-/underweight (savings)") allocated to funds in different atpRating and Morningstar Rating categories, respectively, compared to the empirical distribution of funds in the different categories. In rows "Over-/underweight (investors)", the relative fractions of the total number of investors are shown in relation to the empirical distribution of funds into the different atpRating and Morningstar Rating categories.

Table 8
Average ratings of mutual funds sorted into quartiles as based on inflows to the funds

	Total investments		Number of investors	
Quartile	atpRating	Morningstar	atpRating	Morningstar
1	2.8	3.4	2.9	3.2
2	3.0	2.9	2.9	3.2
3	3.1	3.5	3.1	3.4
4	3.4	3.9	3.4	3.9
4–1	0.54	0.57	0.43	0.76
p-value	2.0%	1.6%	4.2%	0.2%

The table shows the mutual funds sorted into quartiles, as based on the total inflows to the funds from savings in the Special Pension Savings Scheme ("Total investments"), as well as the number of investors having invested in the funds. Quartile 1 has the funds that have received the lowest amounts of inflows (for "Total investments") and the funds that have attracted the smallest numbers of investors (in the columns under "Number of investors"). Quartile 4 has the funds that have attracted the largest amounts of money or the largest number of investors. The table shows the average atpRating and the average Morningstar Rating of the funds in the different quartiles. Row "4–1" shows the differences in average ratings between funds in quartiles 4 and 1, and row "*p*-value" shows *p*-values from one-sided t-tests of whether the differences in row "4–1" are statistically different from zero.

different rating categories. In rows "empirical distribution", we show how the funds are split between the different rating categories, and use this empirical distribution as a benchmark distribution of funds, i.e. the expected allocation of funds, if investors were choosing funds randomly. Disregarding rounding errors, it is seen that around 10% of the funds rated by ATP received the highest and the lowest rating respectively, approximately 22.5% of the funds were in rating category 2 and 4, and approximately 35% of the funds were in rating category 3, as expected. Regarding the funds rated by Morningstar, only 3.42% of funds were in the Morningstar category 1, versus 21.23% in category 5, i.e. also from this table, it can be seen that Danish funds are rated relatively high, as mentioned in footnote 17 and shown in Table 6.

In Table 7, we also show how much more/less savings (investors) went to funds in the different rating categories compared to the empirical distribution of funds. In other words, we look both at the number of investors that choose to invest in funds in different rating categories and the amount of savings that went to the different funds. For instance, 10.98%+4.89%=15.87% of savings went to funds in atpRating category 5, and 19.38% of investors that invested in funds rated by ATP invested in funds in atpRating category 5. This last number corresponds to a 8.40 percentage points overweight of category 5 funds compared to a situation where investors choose funds randomly.

The table reveals that investors have a clear bias in favor of choosing high-rated funds. In particular, 53.6% of savings (54.0% of investors) allocated to funds rated by the atpRating went to funds in atpRating categories 4+5 (the low-cost funds) versus only 23.8% of savings (22.9% of investors) to funds in categories 1+2 (the high-cost funds), i.e. an "overinvestment" in funds in rating-categories 4+5 of 20.08% versus an "underinvestment" of 6.87% in funds in rating categories 1+2. For savings allocated into funds rated by Morningstar, the tendency is even a little stronger: 59.7% of savings (63.7% of investors) went to funds rated 4+5 (the best-performing funds) versus 12.3% of savings (10.7% of investors) to funds rated 1+2 (the worst-performing funds), i.e. an overinvestment in high-rated funds of 12.46% versus an underinvestment of 10.27% in low-rated funds. These numbers constitute clear evidence that

Table 9
Distribution of number of investors relative to number of funds in different rating categories. The first six months of 2005

		ATP								
		1	2	3	4	5	SUM			
MS	1	3.05%	0.00%	−0.50%	−1.04%	−0.63%	0.88%		1+2	4+5
	2	−0.21%	−3.39%	−6.55%	−2.34%	−0.32%	−12.82%	MS	−11.94%	16.47%
	3	2.93%	−4.94%	−1.52%	−3.35%	2.35%	−4.53%	ATP	−1.35%	11.76%
	4	−3.40%	2.21%	−0.66%	−0.68%	0.59%	−1.95%	Both	−0.55%	17.09%
	5	1.30%	1.11%	−1.17%	7.85%	9.33%	18.41%			
	SUM	3.66%	−5.02%	−10.41%	0.44%	11.32%	0.00%			

The table shows the relative number of investors investing in funds ranked by both the atpRating (ATP) and the Morningstar Rating (MS) in relation to the empirical distribution of the number of funds.

high-rated funds get disproportionate high inflows, as also reported for U.S. funds in Del Guercio and Tkac (2001).

We can also test whether investors have statistically significant biases in favor of higher-rated funds. Table 8 shows the average ratings of funds sorted in quartiles according to the amount of savings allocated to the funds (in the columns under "Total Investments") and according to the number of investors having invested in the funds (in the columns under "Number of Investors"). First, note that investors can choose from among 201 different funds, of which 146 are rated by both Morningstar and the atpRating.[20] The numbers in row "Quartile 1", for instance, are thus the average atpRating and Morningstar Ratings for the $146/4 = 36.5 \approx 37$ funds that have received the smallest amounts of inflow (in the columns under "Total Investments") and the numbers in row "Quartile 4" are the average ratings of the $146/4 = 37$ funds that have received the largest amounts of inflow. As can be seen, the average rating for funds that have received the smallest amounts of inflow is 2.8 for the atpRating and 3.4 for the Morningstar Rating, while the corresponding average ratings for funds that have received the largest amounts of inflow are 3.4 for the atpRating and 3.9 for Morningstar. In other words, the average ratings of the funds that have received the largest inflows are higher than the ratings of the funds that have received the least inflow. The differences in average ratings between the funds that have received the lowest amount and the highest amount are 0.54 (atpRating) and 0.57 (Morningstar), and these differences are statistically significantly different from zero, as can be seen in the row "*p*-value". The same pattern of choosing high-rated funds can be observed when considering the number of investors that have invested in the funds as the sorting criteria.

As a robustness check, the same kind of calculations can be made for all 172 funds with an atpRating. The average atpRating for the 25% of the funds that have received the largest amounts of inflow is 3.3, whereas it is only 2.9 for the 25% of the funds that have received the smallest inflow: a difference that is significant at the 1% level.

It could be interesting to see whether investors prefer funds that are rated high by both ratings. Table 9 shows the cross-rating distribution of the number of investors (in percentages) that have invested in the different funds relative to the empirical distribution of the number of funds rated by both ATP and Morningstar. For instance, only 2.05% of the funds that were rated by both rating agencies were "best" in both ratings, i.e. had the lowest costs according to the atpRating and the

[20] This number differs from the number of Danish mutual funds that we have examined so far in this paper. The reason is that only some of the Danish mutual funds are available to Special Pension Savings Scheme investors.

Table 10
Distribution of savings relative to number of funds in different rating categories. The first six months of 2005

		ATP								
		1	2	3	4	5	SUM			
MS	1	4.49%	0.00%	−0.56%	−0.85%	−0.66%	2.42%		1+2	4+5
	2	−0.04%	−3.45%	−6.73%	−2.13%	−0.36%	−12.70%	MS	−10.27%	12.46%
	3	5.07%	−5.12%	−1.59%	−3.42%	2.87%	−2.19%	ATP	1.30%	9.23%
	4	−3.48%	0.71%	0.28%	−1.34%	0.91%	−2.92%	Both	1.01%	13.78%
	5	1.32%	1.80%	−1.94%	7.36%	6.84%	15.39%			
	SUM	7.36%	−6.06%	−10.53%	−0.36%	9.59%	0.00%			

The table shows the relative savings in funds ranked by both the atpRating (ATP) and the Morningstar Rating (MS) in relation to the empirical distribution of the number of funds.

best performance according to Morningstar Rating, but 11.39% of the investors choose to invest in such funds, implying that there were 9.33% more investors that chose the best-rated funds than there were funds rated best by both categories. From the right-hand side of Table 9, it can be seen that there were 17.09% more investors who invested in funds that were assigned categories 4 or 5 by both ratings than there were funds with these two rating categories. Likewise, there were 0.55% less investors in funds in the low-rating categories 1+2 than there were funds in these categories.[21]

Once again, one can evaluate the choices of investors by examining both the number of investors in different funds (as in Table 9), but also by examining the amount of savings in different funds. For completion, Table 10 thus reports the cross-rating distribution of the relative amount of savings in the different funds relative to the empirical distribution of the number of funds rated by both ATP and Morningstar. Again, the results point at a strong investor preference for funds rated high by both ratings. For instance, there were 13.78% more investors investing in funds that were rated high (category 4 or 5) by both ratings than there were funds in these two rating categories, compared with only 1.01% more in funds rated worst (1 or 2) by both rating agencies (the amount invested in the Russian equity fund mentioned in footnote 20 shows up here as well). In other words, whether one looks at the number of investors or the amount of savings, investors have a very clear bias in favor of choosing funds rated high.[22]

Finally, it should be mentioned that when reading Tables 9 and 10, there seem to be some indication that investors distinguish even stronger among funds rated low (or high) by Morningstar than they do for funds rated low (or high) by ATP. For instance, Table 9 shows

[21] Scrutinizing the table, it is seen that there were 3.05% more investors investing in funds with the worst rating in both the atpRating and Morningstar Rating than there were funds rated worst by both rating agencies. The reason for this is one single fund, a Russian equity fund, that during the first six months of 2005 had had an extraordinary high return (indeed, the highest return among all funds). This, to us, suggests two things: (i) investors do not *only* rely on the ratings when choosing funds but also on other information provided by the internet trading platform (even if the results reported in the text show that ratings are indeed very important information to investors), and (ii) when evaluating the impact of ratings, it is more sensible to look at the rating-aggregations "1+2" than focusing on the most extreme ratings, as the few funds in the most extreme rating categories can be strongly influenced by outliers.

[22] The ratings are updated once each quarter. This implies that there is only one rating change with very few funds changing atpRating during the period we look at in this section (January 1, 2005 – June 30, 2005). We leave it to future research, when a sufficiently long history of rating changes is available, to study the consequences of rating changes.

that the overweight for funds rated high (4+5) by Morningstar is 16.47% versus 11.76% for ATP and the underweight is 11.94% for funds rated low by Morningstar versus 1.35% for funds rated low by ATP. One can only speculate why this is so, but given that Morningstar has been present in Denmark for several years, and the atpRating is new, one reason could be that the investor familiarity with the Morningstar Rating is higher than the investor familiarity of the atpRating.[23]

In summary, we find strong evidence that investors seem to prefer investing in high-rated funds. We also find that investors pay attention to both ratings, as they choose funds that have been rated high by both the atpRating and the Morningstar Rating (remember here that Table 6 shows that the ratings contain different information, i.e. the fact that investors choose funds that are rated high by both the atpRating and the Morningstar Rating is not obvious as the two ratings often pick out different funds as the "best performing" funds as explained in Section 6).

Our results line up well with those of Del Guercio and Tkac (2001), who report that US funds receiving four or five stars from Morningstar have strong inflows. In more detail, Del Guercio and Tkac (2001) report that a fund that receives a five-star rating by Morningstar the first time the fund is rated gets a 53% above normal inflow and that an upgrade from being a four-star fund to a five-star fund generates inflows that are 35% higher than normally expected flows. We report a related finding for Danish mutual funds and show that the influence of ratings on investor behavior is not a unique feature of Morningstar Ratings, as the investors we study also choose funds that are rated high by the atpRating.

8. Conclusion

In this paper, we have described and analyzed a cost-based indicator of Danish mutual funds. We use the indicator to sort funds, inspired by the Morningstar Rating, into five different rating categories giving five crowns to the funds with the lowest costs and one crown to the funds with the highest costs.

The development of a cost-based atpRating is based on the idea that costs matter for the after-cost return to the investor, and, at the same time, that costs are considerably more persistent than returns (which are used by, e.g. Morningstar Ratings). Consequently, we test whether the rating of funds, based on their costs, provides information about future risk-adjusted excess returns. We find that the cost rating has some predictive power for returns over a long *out-of-sample* period of eight or ten years. For example, if the investor in 1994 had invested in the 10% of funds that had the lowest costs within their ATP category, she would over the 1994–2003 period have obtained an annual extra 3–4% of risk-adjusted excess returns as compared to investing in the 10% of funds that had the highest costs within their ATP category.

We also find, in accordance with the results of Del Guercio and Tkac (2001), that investors do in fact pay attention to the atpRating, as well as the Morningstar Rating. The data we use here are unique in the sense that we know that the subgroup of investors who have made active

[23] Notice to this, though, that when comparing among all funds rated by ATP versus all funds rated by Morningstar, as in Table 7, this preference for funds rated high by Morningstar is not as clear. A clear difference only appears when comparing among funds rated by both ATP and Morningstar, as in Tables 9 and 10 (remember here that Morningstar requires three years of data before a fund can be rated, whereas ATP only requires cost data for one year before it can be rated, as mentioned in Section 6. For this reason, the number of funds rated by ATP in Tables 9 and 10 is lower than the total number of funds rated by ATP used in Table 7). This could suggest that the atpRating is considered even more important when a fund is not rated by Morningstar.

investment decisions have been exposed to both the Morningstar Rating and the atpRating. Using data on what mutual funds these investors in fact have chosen, we find that investors mainly allocate savings to funds that have been highly rated by the atpRating and/or the Morningstar Rating. Ratings thus seem to have the potential to affect investor decisions.

We have described a cost-based rating of Danish mutual funds. Given that there are indications that the rating predicts the future performance of some Danish mutual funds on long horizons, it would be valuable to evaluate how a cost-based rating performs in a market such as the US market, where there are more mutual funds with long histories on costs and performances. It would also be interesting to study whether changes in the cost rating lead to changes in investors' investment decisions. We leave this for future research.

Appendix A

ATP categories

ATP-category	Morningstar category	Category no.	Mean (%)	Median (%)	Min (%)	Max (%)	SD (%)
I	Denmark — stocks	119	0.99	1.46	−14.17	11.33	4.67
II	Global large cap — stocks	24	0.44	0.87	−13.13	11.61	4.77
II	Global mid/small cap — stocks	25	0.68	0.73	−13.84	14.10	5.02
III	England large cap — stocks	1	0.54	1.10	−10.89	9.15	4.44
III	Europe ex. England — stocks	4	0.83	1.30	−16.08	14.27	5.44
III	Euroland large cap — stocks	6	0.75	1.28	−16.09	12.86	5.43
III	Europe large cap — stocks	9	0.64	1.37	−13.60	11.43	4.90
III	Europe mid cap — stocks	10	0.96	1.35	−18.35	21.83	5.81
III	North America large cap — stocks	13	0.73	0.96	−13.50	13.55	5.35
III	North America mid/small cap — stocks	14	0.98	0.80	−17.77	16.16	6.34
III	Japan large cap — stocks	17	−0.02	−0.09	−13.92	16.38	6.23
III	Japan mid/small cap — stocks	18	0.36	−0.35	−15.64	27.84	7.71
III	Global new markets — stocks	29	0.23	0.67	−28.31	16.92	7.16
IV	DKK short — bonds	121	0.39	0.45	−0.99	1.22	0.39
IV	Euro money market, dynamic	141	0.28	0.25	−0.38	1.18	0.23
V	EUR others — bonds	66	0.44	0.55	−1.67	2.19	0.80
V	DKK others — bonds	120	0.50	0.76	−3.12	2.99	1.00
V	DKK index bonds	142	0.26	0.34	−6.80	6.21	1.68
VI	EUR global — bonds	62	0.36	0.33	−2.30	3.84	1.27
VI	European — bonds	65	0.43	0.49	−2.33	2.87	1.05
VII	England small cap — stocks	3	1.07	1.74	−19.45	17.79	5.84
VII	Europe small cap — stocks	11	0.90	1.26	−19.05	25.72	6.93
VII	Asia ex. Japan — stocks	20	0.22	−0.09	−24.63	20.06	7.38
VII	Asia — stocks	22	0.14	−0.58	−17.45	15.98	6.00
VII	Latin-American — stocks	23	0.63	2.27	−34.97	21.54	8.84
VII	Tech Media Tele — stocks	30	0.93	0.66	−23.78	27.54	8.89
VII	Natural resources — stocks	31	0.26	0.39	−17.67	17.56	5.25
VII	Real estate — stocks	32	0.67	0.64	−8.37	6.35	2.90
VII	Finance — stocks	33	0.88	1.37	−17.74	11.36	5.12
VII	Health — stocks	34	1.01	1.01	−13.11	10.96	5.31
VII	Other sectors — stocks	35	0.51	0.75	−13.06	10.11	4.29
VII	Switzerland — stocks	105	0.65	1.00	−16.22	11.04	4.85
VII	Nordic countries — stocks	107	0.95	1.15	−17.97	20.35	6.21
VII	Central— and Eastern Europe — stocks	110	1.13	1.86	−38.37	24.44	8.58
VIII	EUR Low risk — balanced	36	0.35	0.40	−2.17	2.60	1.04
VIII	EUR moderate risk — balanced	37	0.41	0.75	−6.28	5.83	2.38

(continued on next page)

Appendix A (*continued*)

ATP-category	Morningstar category	Category no.	Mean (%)	Median (%)	Min (%)	Max (%)	SD (%)
VIII	EUR high risk — balanced	38	0.45	0.80	−9.21	7.87	3.38
VIII	USD — balanced	43	0.39	0.79	−7.97	8.70	3.29
VIII	USD — money market	50	0.25	0.01	−6.77	6.66	2.72
VIII	GBP govn. — bonds	56	0.57	0.66	−5.63	6.66	2.38
VIII	EUR high yield — bonds	58	0.22	0.50	−7.46	5.59	1.87
VIII	USD high yield — bonds	60	0.56	0.69	−8.91	8.66	3.42
VIII	New markets — bonds	61	0.82	1.13	−11.70	10.13	3.41
VIII	GBP global — bonds	63	0.47	0.54	−2.71	4.42	1.72
VIII	USD global — bonds	64	0.30	0.08	−4.30	5.17	1.85
VIII	GBP others — bonds	67	0.49	0.54	−5.96	7.09	2.48
VIII	USD others — bonds	68	0.39	0.28	−5.34	7.43	2.55
VIII	Guaranteed funds	69	0.42	0.35	−2.21	2.57	1.00
VIII	NOK others — bonds	109	0.48	0.44	−6.95	7.34	1.90
VIII	SEK others — bonds	118	0.50	0.36	−8.64	8.20	2.22

I: Danish Stocks, II: Global Stocks, III: Regional Stocks, IV: Short Bonds, V: Long Bonds, VI: Global Bonds, VII: Other Stocks, and VIII: Other Bonds. The table shows classifications of Morningstar categories into ATP categories. The table furthermore shows summary statistics (mean, median, maximum, minimum, and standard deviations) of the monthly returns to the funds in the categories.

References

Barber, B.B., Odean, T., Zheng, L., 2005. Out of sight, out of mind: the effects of expenses on mutual fund flows. Journal of Business 78 (6), 2095–2120.

Bechmann, K.L., Rangvid, J., 2007. Danish mutual funds: description, costs, performance, and a European comparison. In: Gregoriou, G.N. (Ed.), Performance of Mutual Funds: An International Perspective. Palgrave–MacMillan. Chapter 3.

Blake, C.R., Morey, M.R., 2000. Morningstar Ratings and mutual fund performance. Journal of Financial and Quantitative Analysis 35 (3), 451–483.

Blake, D., Timmermann, A., 1998. Mutual fund performance: evidence from the UK. European Finance Review 2 (1), 57–77.

Blake, C.R., Elton, E.J., Gruber, M.J., 1993. The performance of bond mutual funds. Journal of Business 66 (3), 371–403.

Cai, J., Chan, K.C., Yamada, T., 1997. The performance of Japanese mutual funds. Review of Financial Studies 10 (2), 237–273.

Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of Finance 52 (1), 57–82.

Cesari, R., Panetta, F., 2002. The performance of Italian equity funds. Journal of Banking and Finance 26 (1), 99–126.

Chalmers, J.M.R., Edelen, R.M., Kadlec, G.B., 1999. Transaction-costs expenditures and the relative performance of mutual funds. Manuscript. The Wharton Financial Institutions Center.

Choi, J.J., Laibson, D., Madrian, B.C., 2006. Why does the law of one price fail? An experiment on index mutual funds. NBER WP No. 12261.

Christensen, M., 2003. Evaluating Danish mutual fund performance. Manuscript. Aarhus School of Business.

Christensen, M., 2005. Danish mutual fund performance – selectivity, market timing and persistence. Working Paper F-2005-01, Finance Research Group. Aarhus School of Business.

Cronqvist, H., 2006. Advertising and portfolio choice. Manuscript. Fisher College of Business, Ohio State University.

Cronqvist, H., Thaler, R.H., 2004. Design choices in privatized social-security systems: learning from the Swedish experience. American Economic Review 94 (2), 424–428.

Dahlquist, M., Engström, S., Söderlind, P., 2000. Performance and characteristics of Swedish mutual funds. Journal of Financial and Quantitative Analysis 35 (3), 409–423.

Deaves, R., 2004. Data-conditioning biases, performance, persistence and flows: the case of Canadian equity funds. Journal of Banking and Finance 28 (3), 673–694.

Del Guercio, D., Tkac, P.A., 2001. Star Power: the effect of Morningstar Ratings on mutual fund flows. Working Paper 2001–15. Federal Reserve Bank of Atlanta.

Elton, E.E., Gruber, M.J., Das, S., Hlavka, M., 1993. Efficiency with costly information: a reinterpretation of evidence from managed portfolios. Review of Financial Studies 6 (1), 1–22.
Gruber, M.J., 1996. Another puzzle: the growth in actively managed mutual funds. Journal of Finance 51 (2), 783–810.
Holland, P.W., Welsch, R.E., 1977. Robust regression using iteratively reweighted least-squares. Communications in Statistics: Theory and Methods A6, 813–827.
Karlsson, A., Massa, M., Simonov, A., 2006. Portfolio choice and menu exposure. Manuscript. Stockholm School of Economics.
Korkeamaki, T.P., Smythe Jr., T.I., 2004. Effects of market segmentation and bank concentration on mutual fund expenses and returns: evidence from Finland. European Financial Management 10 (3), 413–438.
Malkiel, B.G., 1995. Returns from investing in equity mutual funds 1971 to 1991. Journal of Finance 50 (2), 549–572.
Morey, M.R., 2002. Rating the raters: an investigation of mutual fund rating services. Journal of Investment Consulting 5 (2), 30–50.
Otten, R., Bams, D., 2002. European mutual fund performance. European Financial Management 8 (1), 75–101.
Vinod, H.D., Morey, M.R., 2002. Estimation risk in Morningstar mutual fund ratings. Journal of Investing 11 (4), 67–75.
Wermers, R., 2000. Mutual fund performance: an empirical decomposition into stock-picking talent, style, transactions costs, and expenses. Journal of Finance 55 (3), 1655–1695.