Why you should care about investment costs: A risk-adjusted utility approach

Morten Tolver Kronborg[⋆†] and Søren Fiig Jarner[⋆]

January 15, 2015

This is an accepted manuscript of an article published by Elsevier in Journal of Behavioral and Experimental Finance on 28/03/2015, available online at http://www.sciencedirect.com/ (DOI: 10.1016/j.jbef.2015.03.004)

Abstract

Under the assumption of zero correlation between cost ratios and expected investment returns we analyze the impact of proportional investment costs. We consider a constant relative risk aversion investor optimizing expected utility from terminal wealth and identify, in addition to the direct effect due to the additional costs incurred, an indirect effect. The indirect effect is due to lost investment opportunities and a less risky stock position induced by investment costs. By use of an indifferent compensation measure, defined as the minimum relative increase in the initial wealth the investor demands in compensation to accept incurring investment costs of a certain size, we quantify the impact of investment costs. We obtain for realistic parameters that the indirect effect is between half and the same size as the direct effect, and that the investment decision seems to be of very little importance compared to the size of the investment costs.

JEL classification: C61.

Keywords: Investment costs; risk aversion; indifferent compensation measure; certainty equivalent.

[⋆]Authors at: ATP (Danish Labour Market Supplementary Pension Scheme), Kongens Vænge 8, 3400 Hillerød, Denmark, and University of Copenhagen, Department of Mathematical Sciences, Universitetsparken 5, 2100 Copenhagen Ø, Denmark.

† Corresponding author. E-mail address: mtk@atp.dk, Tel.: +45 48204537.

1 Introduction

Most investors seem primarily to focus on the ability of excellent stock picking, when deciding which fund should manage their savings. At the same time costs charged by funds seem to differ by a great deal. The average U.S. equity mutual fund charges around 1.3–1.5 percent, but cost ratios range from as low as 0.2 (index funds) to as high as 2 percent. In general, costs can vary substantially across comparable funds, and larger funds and fund complexes charge lower costs (see e.g. Khorana et al. (2008)). Clearly, the argument for charging high costs is excellent stock picking. The managers of expensive funds are likely to claim that the additional return they are expected to generate (compared to any cheaper fund manager) more than compensates for the extra costs. However, the vast majority of the large literature finds that higher costs are not related to superior returns (see e.g. Gil-Bazo and Ruiz-Verdú (2009), Carhart (1997), Fama and French (2010), Malkiel (1995) and Malhotra and Mcleod (1997)).

In particular this is demonstrated by Gil-Bazo and Ruiz-Verdú (2009) who consider a data set including all open-end U.S. mutual funds that were active in the 1961 to 2005 period. They consider a series of robustness checks consisting of checking for the impact of funds with extreme cost ratios and extreme risk-adjusted performance; the impact of small funds; exclusive focusing on funds for which annual operating costs account for 100 percent of all costs or focusing only on funds with loads; splitting time into subperiods; splitting mutual funds into categories. In all cases the conclusion of Gil-Bazo and Ruiz-Verdú (2009) stays the same: The hypothesis of a unit slope relation between risk-adjusted before-fee performance and cost ratios falls at any conventional significance level. In fact the expected "additional" before-fee return is, ironically, estimated by Gil-Bazo and Ruiz-Verdú (2009) to -0.63% per 1% increase of the cost ratio. In relation Carhart (1997), who using the same data set, concludes that higher costs depress investment performance while increasing fund companies' profitability. Also Fama and French (2010) report that only very few funds produce benchmark-adjusted expected returns sufficient to cover their costs. One of the reasons that some funds are more expensive is due to the more actively managed investments. Huang et al. (2013) report, using a sample of 2979 U.S. equity funds over the period between 1980 and 2009, that the top and bottom decile of funds on average change their annualized volatility by more than six percentage points. They also find, by use of a holding-based measure of risk shifting, that funds which alter risk perform worse than funds that keep stable risk levels over time, suggesting that risk shifting either is an indication of inferior ability or is motivated by agency issues. Summing up, it seems hard to prove that good performance is anything but a random phenomena.

Consequently, we analyze the impact of investment costs under the assumption of a zero correlation between the cost ratio and the expected investment return. However, note that the analysis also applies to the situation where we assume that funds can indeed generate (some) excess return, in which case, the cost ratio should be interpreted as the net-cost. The literature, e.g. the references above, seems only to focus on the loss in rate of return. However, the loss in rate of return simply induced by paying higher investment costs might not describe the actual loss suffered by the investor. A more sophisticated approach would be to take into account the risk aversion of the investor when evaluating the impact of investment costs, thereby also introducing a change in the investment strategy induced by investment costs. Introducing proportional investment costs and by use of utility functions, this is the approach taken in our paper. Two related papers, also taking the investor's risk aversion into account while considering proportional costs are Guillén et al. (2014) and Palczewski et al. (2013) (the latter analyzes the impact of transaction costs).

Guillén et al. (2014) consider a Value at Risk investor (VaR-investor) who invests in a Black-Scholes market concerned about a given α-percent quantile of the terminal wealth distribution. By introducing investment costs the investor is forced to invest less in the stock market in order to maintain the same α-percent quantile. Consequently, the loss in the geometric rate of return splits into two effects: (a) A direct effect due to the additional expense incurred and (b) an indirect effect due to a less risky stock position. Some of the capital the investor, prior to introducing investment costs, was willing to risk losing is now used to pay investment costs. The main drawback of the VaR-approach is that no monetary quantification of how much the investor actually suffers from investment costs seems to be possible. Focusing at the geometric rate of return seems a bit ad hoc since, in the first place, when deciding upon the investment strategy, the VaR-investor had no particular preferences for a high median. Using the geometric rate of return to measure the impact of investment costs also restricts the parameter space since for very risk seeking investors, introducing investment costs actually *increases* the geometric rate of return.

In contrast, Palczewski et al. (2013) use utility functions, but focus instead on the impact of transaction costs. They optimize expected utility from investing in a market consisting of a risk free asset and a risky asset modeled by a diffusion model with state-dependent drift. The effect of costs can again be divided into a direct and an indirect effect. This time the indirect effect is due to less trading in the asset portfolio. By calculating the indifference price, defined as the amount of money the investor is willing to pay up front to avoid incurring transaction costs, they find that in general the loss in utility due to proportional transaction costs is about twice as large as the direct expenses incurred.

Similar findings are offered by our paper for the case of proportional investment costs. We focus on a constant relative risk aversion (CRRA) utility optimizer who hands over his savings to a fund investing in a frictionless Black-Scholes market while being charged proportional investment costs. In contrast to the VaR-approach of Guillén et al. (2014) the change in investment strategy and, consequently, the change in geometric rate of return induced by a change in investment costs becomes independent of the investment horizon. The change in geometric rate of return is the same for both short and long term investors. As in Guillén et al. (2014) and Palczewski et al. (2013) we obtain a direct and an indirect effect of costs. In our case, the indirect effect is the sum of 1) lost investment opportunities, since the amount of money available for investment is reduced, and 2) the effect from a changed asset allocation induced by the change in costs. In order to quantify the financial impact of investment costs we calculate the indifferent compensation ratio (ICR), defined as the minimum relative increase in the initial wealth the investor demands in compensation to accept incurring investment costs of a certain size. For a CRRA utility optimizing investor the ICR is proved to be equal to the relative change in certainty equivalents. By comparing the ICR value to the financial value of accumulated investment costs, we find, similar to Palczewski et al. (2013), that the magnitude of the indirect effect exceeds the direct effect when considering a long-term investor (40 years horizon, i.e. investing for retirement). That is, the amount of money needed up front to be compensated for investment costs can be twice as big as the financial value of accumulated investment costs, i.e. the amount of money needed to replicate the cost expenses. For a short term investor we find that the magnitude of the indirect effect is half the size of the direct effect. In the words of Jens Perch Nielsen, this can be summarized by the catchy phrase: *The double blow of investment costs*. Finally, we undertake a study of whether the investment strategy or the size of investment costs is of most importance. Specifically, we study an investor facing high investment costs and an optimal investment strategy (w.r.t. his risk aversion profile) and ask which suboptimal investment strategies the investor is willing to accept if he at the same time is offered lower investment costs. The conclusion is independent of the time horizon and very clear: The asset allocation is of very little importance compared to the size of investment costs.

The analysis is performed for a CRRA utility optimizing investor paying proportional investment costs. We have deliberately chosen CRRA utility and a simple fee model to highlight the points we wish to make without obscuring the analysis with technicalities. More complex fee structures can also be analyzed, see e.g. Janeček and Sîrbu (2012), at the expense of more technicalities and less explicit solutions.

The outline of the paper is as follows: In Section 2 we introduce the investor and the financial market to be considered, the wealth dynamics, and quantiles for the terminal wealth distribution together with the geometric rate of return. In Section 3 we analyze for a utility optimizing investor the change in investment strategy induced by a change in investment costs, and we compare the results with the VaR-approach by Guillén et al. (2014). In Section 4, we use the indifferent compensation measure to quantify the impact of investment costs, and to evaluate whether the investment strategy or the size of investment costs is of most importance. Finally, Section 5 summarizes the main conclusions.

2 The financial model

Consider an investor with time horizon $T > 0$ who has the possibility to invest in a Black-Scholes market given by a risky stock, S, and a risk-free bank account, B, with dynamics given by

$$\begin{aligned}
dB(t) &= rB(t)dt,\ B(0) = 1, \\
dS(t) &= \mu S(t)dt + \sigma S(t)dW(t),\ S(0) = s_0 > 0.
\end{aligned}$$

We assume the risk-free rate, r, the expected return of the stock, μ, and the volatility of the stock, σ, to be constant with $\mu > r$. The risky part of the stock, W, is a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$ equipped with the filtration $\mathbb{F}^W = (\mathcal{F}^W(t))_{t\in[0,T]}$ given by the P-augmentation of the filtration $\sigma\{W(s);\, 0 \le s \le t\}, \forall t \in [0, T]$. In principle, a financial market model should contain more than one stock. However, it is well known from the Mutual Fund Theorem by Merton (1971) that for a HARA (Hyperbolic Absolute Risk Aversion) utility investor, as we shall consider, the optimal asset allocation will always be between an optimal portfolio of the stocks weighted by their Sharp ratios, and the risk-free bank account (at least for diffusion processes with deterministic coefficients).

Let X denote the wealth process of the investor and π the proportion of the wealth invested in risky stocks. Consequently, the proportion $1 - \pi$ of the wealth is invested in the risk-free bank account. Assume further that the investor hands over the asset management to a second party (e.g. a mutual fund or a pension company), and therefore is subject to investment costs. More precisely, assume that a constant fraction, ν, of the amount of money invested in the stock market is deducted from the wealth process. The dynamics of the wealth process becomes

$$\begin{aligned}
dX(t) &= \left((1-\pi(t))\frac{dB(t)}{B(t)} + \pi(t)\frac{dS(t)}{S(t)} - \pi(t)\nu\right) X(t) \\
&= (r + \pi(t)(\mu - \nu - r))X(t)dt + \sigma\pi(t)X(t)dW(t), \qquad (1) \\
X(0) &= x_0.
\end{aligned}$$

Note that costs from investing in the risk-free asset are not explicitly indicated, but are indirectly present through a possibly lower risk-free rate of return. However, compared to costs emerging from investing in stocks, costs from investing in the risk-free bank asset are likely to be negligible. Moreover, we assume through out the paper that the expected excess return is greater than investment costs, $\mu - r > \nu$. Thereby, the net expected return is assumed positive such that there is an incentive to buy stocks.

2.1 Wealth quantiles for a constant investment strategy

One extremely popular investment strategy is to hold, at all times, a constant fraction of wealth in stocks. This also turns out to be optimal for the popular CRRA (Constant Relative Risk Aversion) utility optimizing investor introduced by the pioneering work of Merton (1969). In this paper we consider only such constant strategies. Note that the word *constant* is rather misleading since the amount of money invested in stocks is dynamically reallocated to keep the fraction of wealth invested in stocks constant. Letting π denote a constant investment strategy, the wealth dynamics (1) takes the form of a geometric Brownian motion with solution

$$X(t) = x_0 \exp\left(\left(r + \pi(\mu - \nu - r) - \frac{1}{2}\pi^2\sigma^2\right)t + \pi\sigma W(t)\right). \qquad (2)$$

Since $W(t) \stackrel{\mathcal{D}}{=} \sqrt{t}U$, where U is standard normal distributed, we get from (2) that the α-quantile, q_α, of the terminal wealth distribution is given by

$$q_\alpha(\nu,\pi) = x_0 \exp\left(\left(r + \pi(\mu - \nu - r) - \frac{1}{2}\pi^2\sigma^2\right)T + \pi\sigma\sqrt{T}d_\alpha\right), \tag{3}$$

where d_α is the α-quantile of the standard normal distribution. In particular, the median is given by

$$q_{50}(\nu,\pi) = x_0 \exp\left(\left(r + \pi(\mu - \nu - r) - \frac{1}{2}\pi^2\sigma^2\right)T\right).$$

From this we get that the median rate of return (geometric rate of return), from now on referred to as simply *the rate of return*, ρ, is given by

$$\rho(\nu,\pi) = r + \pi(\mu - \nu - r) - \frac{1}{2}\pi^2\sigma^2. \tag{4}$$

Note that the rate of return, in contrast to the expected arithmetic rate of return, possesses a maximum with respect to the risky stock allocation.

3 Investment costs' impact on the investment strategy

Consider a *naive investor* who for some reasons invests, without any objective in mind, a constant fraction, π, of his wealth in risky stocks. For the naive investor the gain/loss in terms of rate of return caused by a change in the investment costs from ν_1 to ν_2 is given by

$$\rho(\nu_2,\pi) - \rho(\nu_1,\pi) = \pi(\nu_1 - \nu_2), \tag{5}$$

where ρ is defined by (4). This simple calculation seems to be how investment costs are normally quantified.

Remark 3.1. *As stated in the Introduction, a large part of the literature finds that higher costs are not related to superior returns. Our analysis and interpretations are based on this assumption. However, from the point of view of the investor, who thinks that higher costs imply higher expected returns, the parametrization could be interpreted as* $\nu_1 = \widetilde{\nu_1} + (\mu - \widetilde{\mu_1})$*, where* $\widetilde{\nu_1}$ *and* $\widetilde{\mu_1}$ *are the actual investment costs and the actual (larger) expected return, respectively. In other words, the cost ratio could be interpreted as the net-cost, and the analysis applies also to this case.*

However, a *sophisticated investor* should take his risk aversion into account and thereby adjust his investment strategy accordingly when investment costs change. Our main focus will be how a utility investor adjusts his investment strategy in response to changes in the investment costs and how to quantify the associated gain or loss. However, for comparison, we first consider a VaR-investor as done by Guillén et al. (2014). The main point is that the sophisticated investor experiences a greater change in the rate of return induced by a change in investment costs compared to the naive investor.

3.1 Base case parameter values

Throughout this paper we will, as a base case example for illustrating results, use the market parameters $\mu = 7\%$, $\sigma = 20\%$ and $r = 3\%$. When illustrating the impact of investment costs we consider the difference between paying investment costs at rate $\nu_1 = 1.4\%$ and $\nu_2 = 0.6\%$, respectively. Most papers focus on the impact of introducing investment costs. We focus on an individual who prefers to hand over the asset management to a fund. In that situation it seems unattainable to obtain zero investment costs. Therefore, we consider instead the "price" of investing through an expensive fund (ν_1) compared to a cheap fund (ν_2). The investment costs considered are consistent with those observed in the market (see references in Section 1).

Later, for robustness check and to gain further insight, we present results for varying values of the market parameters, the risk aversion and horizon of the investor, and the cost ratios.

3.2 The Value at Risk approach

This subsection briefly presents and explains the impact of investment costs for a VaR-investor, as presented in Guillén et al. (2014). A VaR(α)-investor is an investor being concerned with the value of the α-quantile of the terminal wealth distribution, i.e. with the level for which the realized terminal wealth risk to fall below with a probability of α-percent. Denote the lower level of wealth by $\widehat{X}(T)$. To be more precise, the VaR(α)-investor facing investment costs ν_1 chooses his investment strategy, $\widehat{\pi}_1$, such that

$$q_\alpha(\nu_1, \widehat{\pi}_1) = \widehat{X}(T),$$

where q_α is given by (3). The point is that if investment costs change from ν_1 to ν_2 this relation holds no more. Obviously, the VaR-(α)-investor should adjust his investment strategy such that the VaR-criteria is still fulfilled. We therefore look for a change in the investment strategy, Δ, fulfilling the relation

$$q_\alpha(\nu_2, \widehat{\pi}_1 + \Delta) = q_\alpha(\nu_1, \widehat{\pi}_1), \tag{6}$$

i.e. being faced with investment costs ν_2 the investor should instead invest $\widehat{\pi}_2 = \widehat{\pi}_1 + \Delta$ of his wealth in risky stocks. We get the solution

$$\Delta = \frac{-b \pm \sqrt{D}}{2a},$$

where $D = b^2 - 4ac$ and

$$\begin{aligned} a &= \frac{\sigma}{2}, \\ b &= -(\mu - \nu_2 - r) + \widehat{\pi}_1\sigma^2 - \frac{\sigma d_\alpha}{\sqrt{T}}, \\ c &= \widehat{\pi}_1(\nu_2 - \nu_1). \end{aligned}$$

Of the two solutions we focus on the one allowing for more stocks in the case of lower costs ($\nu_2 < \nu_1$) and fewer stocks in the case of higher costs ($\nu_2 > \nu_1$). Guillén et al. (2014) derive restrictions on the parameter space ensuring that the median increases and the α-quantile decreases when exposure in the risky asset increases. In contrast to (5) the change in rate of return becomes

$$\rho(\nu_2, \widehat{\pi}_2) - \rho(\nu_1, \widehat{\pi}_1) = \widehat{\pi}_1(\nu_1 - \nu_2) + \Delta(\mu - \nu_2 - r) - \frac{1}{2}\Delta^2\sigma^2 - \widehat{\pi}_1\Delta\sigma^2. \tag{7}$$

Note that since Δ and $\widehat{\pi}_1$ depend on the investment horizon so does the change in rate of return. Figure 1 illustrates for the base case example (see Subsection 3.1) the VaR-calibration concept for an investor investing one unit initially. First, when investment costs are $\nu_1 = 1.4\%$ the VaR(10%)-investor chooses to invest $\widehat{\pi}_1 = 60\%$ of his wealth in stocks corresponding to a 10-percent quantile for the terminal wealth distribution of $q_{10\%}(\nu_1, 60\%) = 1.75$ and a median of $q_{50\%}(\nu_1, 60\%) = 4.65$. Not changing the investment strategy, but now being charged $\nu_2 = 0.6\%$ in investment costs, the 10-percent quantile becomes $q_{10\%}(\nu_2, 60\%) = 2.12$ and the median increases, due to paying less costs, to $q_{50\%}(\nu_2, 60\%) = 5.63$. However, the investor is able to increase the fraction of wealth invested in stocks to $\widehat{\pi}_2 = 74.4\%$ thereby obtaining the original 10-percent target quantile $q_{10\%}(\nu_2, 74.4\%) = q_{10\%}(\nu_1, 60\%) = 1.75$ and at the same time an even bigger median of size $q_{50\%}(\nu_2, 74.4\%) = 5.87$. In other words, Figure 1 illustrates a direct and an indirect effect of investment costs.

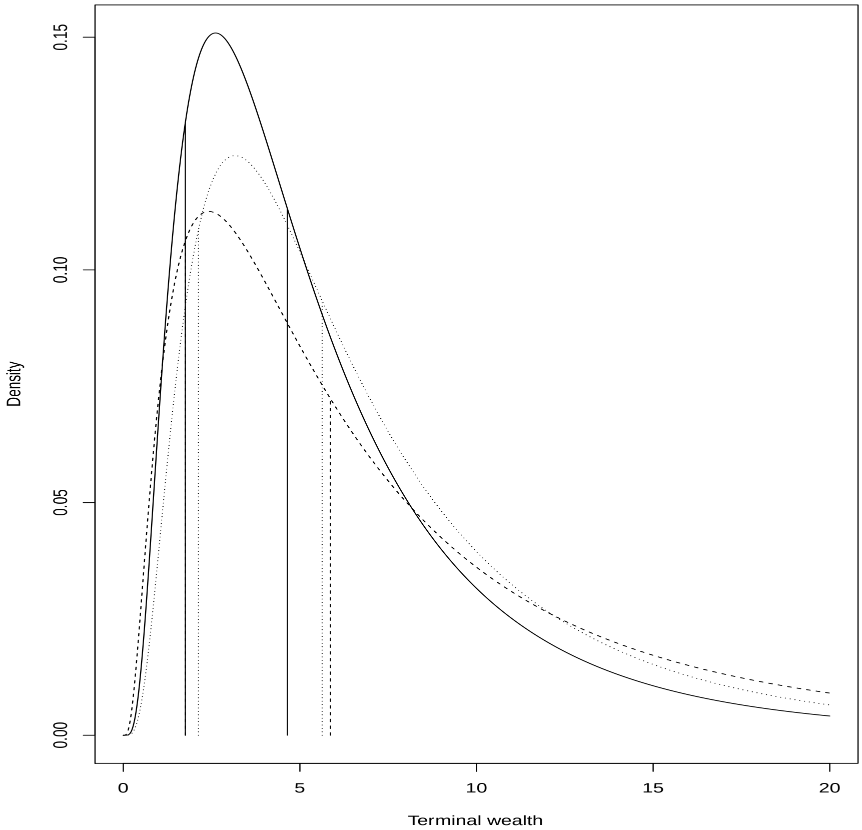

Figure 1: Illustration of the VaR-calibration: Terminal wealth distribution for an investor with time horizon $T = 40$ years investing $\widehat{\pi}_1 = 60\%$ (solid curve), $\pi_2 = 60\%$ (dotted curve) or $\widehat{\pi}_2 = 74.4\%$ (dashed curve) of his wealth in stocks while being charged $\nu_1 = 1.4\%$, $\nu_2 = 0.6\%$ or $\nu_2 = 0.6\%$ in investment costs, respectively. The 10-percent quantiles and the medians are indicated by vertical lines. Initial wealth is one monetary unit.

Guillén et al. (2014) focus on the loss in rate of return incurred by the VaR-investor from introducing investment costs. For some parameters, going from no costs to costs actually charged by funds around the world, they conclude that the loss in rate of return calculated by (7) is double the size of the naive loss calculated by (5), i.e. an indirect effect equal to the direct effect. However, going from low investment costs (say) 0.6 percent to higher costs at (say) 1.4 percent, both within the interval of normal costs charged by real life funds, we get a much more modest additional loss in the rate of return; the loss is about 1/5 higher than the loss from the naive calculation. To be concrete, for the case illustrated in Figure 1, we obtain a direct loss in the rate of return of 0.48 percent, and an indirect loss of 0.10 percent.

One question the VaR-approach is not capable of answering is how to evaluate the impact of investment cost taking the entire distribution of terminal wealth into account. E.g. how much is it “worth” being charged ν_2 percent instead of ν_1 percent? This is the reason we take on another approach and consider a utility optimizing investor instead.

3.3 The utility approach

Consider an investor measuring his attitude towards risk by a constant relative risk aversion (CRRA) function defined by

$$u(x) = \begin{cases} \frac{x^{1-\gamma}}{1-\gamma}, & \text{if } x > 0, \\ \lim_{x \searrow 0} \frac{x^{1-\gamma}}{1-\gamma}, & \text{if } x = 0, \\ -\infty, & \text{if } x < 0, \end{cases} \tag{8}$$

for some $\gamma \in (0, \infty) \backslash \{1\}$ ($\gamma = 1$ corresponds to the logarithmic case). A commonly used object for the investor is to try maximizing the expected utility from terminal wealth. Formally, the investor looks for the investment strategy π^* fulfilling

$$\sup_{\pi} E\left[u(X^{\pi}(T))\right] = E\left[u\left(X^{\pi^*}(T)\right)\right], \tag{9}$$

where we have indicated the π-dependence of the terminal wealth in the notation. The problem given by (9) is one of the most well-known optimization problems from the literature of financial stochastic control and is often referred to as *Merton's problem* after the pioneering work of Merton (1969). The solution to the problem can be obtained by use of dynamic programming, which turns the stochastic optimization problem into a deterministic optimization problem and a set of partial differential equations (Hamilton-Jacobi-Bellman equations), or by use of the more direct method called the "Martingale method" developed by Karatzas et al. (1987) and Cox and Huang (1989), which turns the problem into a static optimization problem and a representation problem. For a nice introduction to financial stochastic control and to see several approaches to solving Merton's problem (9) see Rogers (2013). The solution turns out to depend on the investor's risk aversion, the riskiness of stocks (volatility) and the stock risk premium net investment costs, and is given by

$$\pi^*(\nu) = \frac{1}{\gamma} \frac{\mu - \nu - r}{\sigma^2}. \tag{10}$$

We see that the optimal investment strategy is to hold a constant fraction of wealth in risky assets. This is a feature of CRRA utility which, from a mathematical point of view, is rather appealing (see Subsection 2.1). The strategy itself dictates directly the change in the risky position induced by a change in investment costs. We get

$$\Delta = \pi^*(\nu_2) - \pi^*(\nu_1) = \frac{1}{\gamma} \frac{\nu_1 - \nu_2}{\sigma^2}.$$

As for the VaR-approach we see that, naturally, if the investment costs decrease (increase) more stocks should be bought (sold). However, opposed to the VaR-approach the change in the risky position is always unique (no short-selling solution). Even more appealing we get in contrast to the VaR-approach (see (7)) that the change in rate of return becomes independent of the investment horizon. We get, with $\pi_i^* = \pi^*(\nu_i)$, that

$$\rho(\nu_2, \pi_2^*) - \rho(\nu_1, \pi_1^*) = \pi_1^*(\nu_1 - \nu_2) + \Delta(\mu - \nu_2 - r) - \frac{1}{2}\Delta^2\sigma^2 - \pi_1^*\Delta\sigma^2.$$

Clearly, the rather large change in rate of return induced by a change in investment costs, was for the VaR-investor partly because a very long investment horizon was considered. Since many private investors mainly invest through being a member of a pension scheme, i.e. long time horizon, the conclusion still seems to be highly relevant. However, the fact that for the utility investor the change in rate of return induced by a change in investment costs does not depend on the investment horizon counters criticism claiming that the effect is negligible for short horizons.

Figure 2 illustrates for the base case example (see Subsection 3.1) the utility-calibration concept for an investor with risk aversion corresponding to $\gamma = 1.0833$, who invests one unit initially. First, when investment costs are $\nu_1 = 1.4\%$ the investor optimally invests $\pi_1^* = 60\%$ of his wealth in stocks. When investment costs decline to $\nu_2 = 0.6\%$ the investor increases this fraction to $\pi_2^* = 78.5\%$. In contrast to the VaR-approach we see in Figure 2 that the 10-percent quantiles of the terminal wealth distributions are not equal for the two cases. The utility investor does not focus solely on the 10-percent quantile, but evaluates the entire distribution of terminal wealth when he decides how much of the saved investment costs to invest in stocks. In this respect he acts more sophisticated than the VaR-investor.

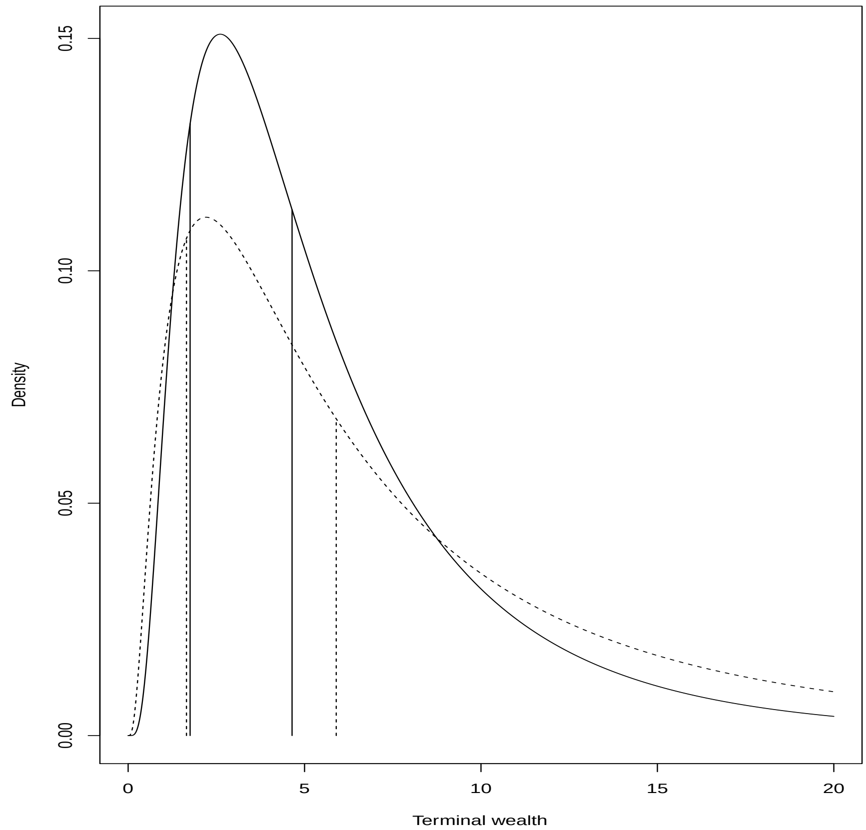

Figure 2: Illustration of the utility-calibration: Terminal wealth distribution for an investor with time horizon $T = 40$ years investing $\pi_1^* = 60\%$ (solid curve) or $\pi_2^* = 78.5\%$ (dashed curve) of his wealth in stocks while being charged $\nu_1 = 1.4\%$ or $\nu_2 = 0.6\%$ in investment costs, respectively. The 10-percent quantiles and the medians are indicated by vertical lines. Initial wealth is one monetary unit.

3.4 Comparing the VaR and utility approach

In Subsections 3.2 and 3.3 we considered a VaR-investor and a utility optimizing investor, respectively. The VaR-investor picked his investment strategy in order to obtain a target 10-percent quantile for the terminal wealth distribution equal to 1.75. The utility investor picked his investment strategy in order to maximize expected utility from terminal wealth while using the power utility function given by (8) with risk aversion given by $\gamma = 1.0833$. In order to be able to compare the two investors' reaction to changes in investment costs we have constructed the two examples such that when investment costs are $\nu_1 = 1.4\%$ both investors prefer to invest $\pi = 60\%$ of the wealth in stocks.

Changing investment costs from $\nu_1 = 1.4\%$ to $\nu_2 = 0.6\%$ we saw in Subsections 3.2 and 3.3 that the change in the investment strategy was $\Delta = 14.4\%$ for the VaR-investor and $\Delta = 18.5\%$ for the utility investor. From this example we conclude that the utility investor reacts more strongly compared to the VaR-investor when investment costs change. However, as illustrated by Figure 3, this is not a general rule of thumb. In fact, by changing the risk profiles such that both investors prefer a $\pi = 30\%$ position of wealth in stocks when investment costs are $\nu_1 = 1.4\%$, lowering the investment costs now makes the VaR-investor change his investment strategy more than the utility optimizing investor.

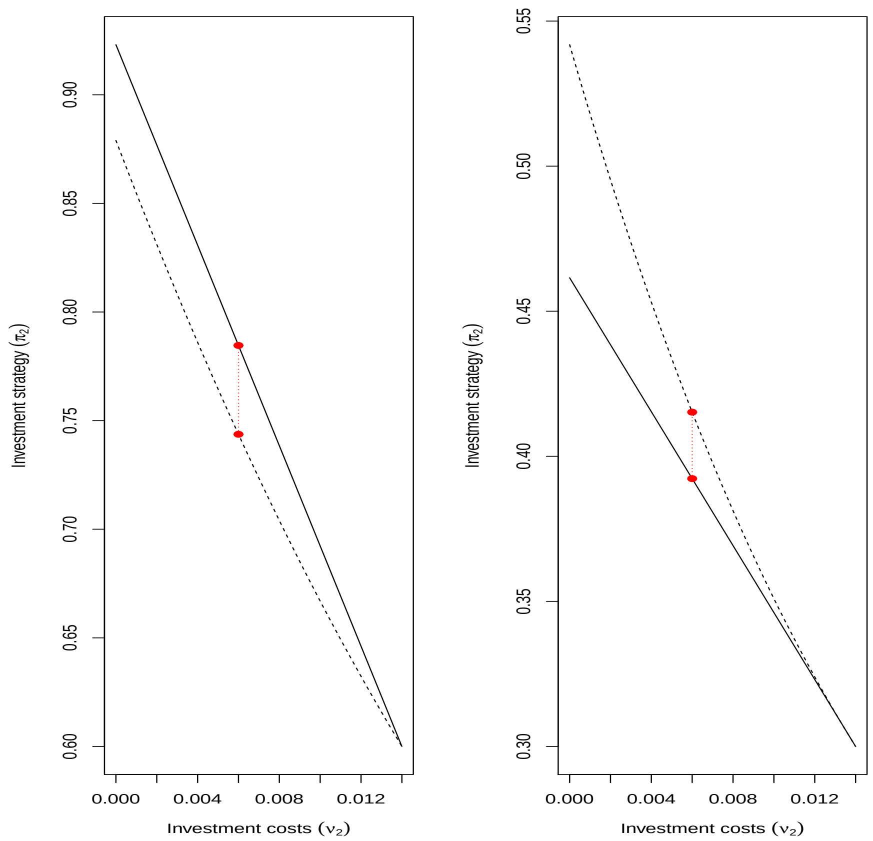

Figure 3: Illustrations of how the utility optimizing investor (solid curves) and the VaR-investor (dashed curves) change investment strategy when investment costs decrease from $\nu_1 = 1.4\%$ to $\nu_2 \in [0, \nu_1]$. The left plot illustrates the case where the investors prefer a $\pi = 60\%$ allocation to stocks, and the right plot the case of a preferable $\pi = 30\%$ allocation to stocks, when investment costs are $\nu_1 = 1.4\%$. The dots indicate the base case example where investment costs decrease to $\nu_2 = 0.6\%$. The investment horizon is $T = 40$ years.

Basak and Shapiro (2001) finds the optimal investment strategy for a power utility investor faced with a VaR restriction on terminal wealth. More precisely, they optimize expected utility from terminal wealth under the constraint that terminal wealth is allowed to fall below a certain lower bound with probability at most α. It is shown that the optimal strategy consists of an optimal unrestricted portfolio and an insurance covering only intermediate losses in that portfolio, i.e. the worst losses in the terminal wealth distribution are accepted. The investor considered by Basak and Shapiro (2001) is motivated by the widespread use of VaR-based risk management among institutional investors, either voluntarily or enforced by regulation. This investor is in a sense a combination of the two investors we consider; in response to changed investment costs the investor of Basak and Shapiro (2001) would change investment strategy to optimize utility *and* adhere to a VaR-constraint, while the investors we consider would focus on one of these measures only. The results of Basak and Shapiro (2001) offer an alternative calibration of the investment strategy and thereby an alternative way to quantify the effect of investment costs. However, our aim is to quantity the effect of investment costs for private investors for which a combined utility and VaR-approach does not appear natural. Furthermore, we prefer the simpler model of Subsection 3.3 in order to obtain closed-form results and clear interpretations when examining the impact of investment costs.

4 Quantification of the impact of investment costs

In this section we use the utility approach described in Subsection 3.3 to quantify the impact of investment costs. Inserting the optimal investment strategy (10) into (2) we are able to calculate the optimal expected utility given by (9). We get

$$\begin{aligned} &\sup_{\pi} E\left[u\left(X^{(\nu,\pi)}(T)\right)\right] \\ &= E\left[u\left(X^{(\nu,\pi^*)}(T)\right)\right] \\ &= \frac{1}{1-\gamma} x_0^{1-\gamma} \exp\left(\left(r + \pi^*(\mu - r - \nu) - \frac{1}{2}\left(\pi^*\right)^2 \sigma^2\right) T(1-\gamma) + \frac{1}{2}\left(\pi^*\right)^2 \sigma^2 (1-\gamma)^2 T\right) \\ &= \frac{1}{1-\gamma} x_0^{1-\gamma} \exp\left\{\left(r + \frac{1}{2}\frac{1-\gamma}{\gamma}\left(\frac{\mu - r - \nu}{\sigma}\right)^2\right) T\right\}, \qquad (11) \end{aligned}$$

where x_0 denotes the initial wealth. Now, define the Indifferent Compensation Ratio as the minimum relative increase in the initial wealth the investor demands in compensation to accept incurring higher investment costs. Formally, the indifferent compensation ratio for two given levels of investment costs, $ICR_{(\nu_1,\nu_2)}$, is given by the relation

$$\begin{aligned} &\sup_{\pi} E\left[u\left(X^{(\nu_2,\pi)}(T)\right) \middle| X^{(\nu_2,\pi)}(0) = x_0\right] \\ &= \sup_{\pi} E\left[u\left(X^{(\nu_1,\pi)}(T)\right) \middle| X^{(\nu_1,\pi)}(0) = x_0\left(1 + ICR_{(\nu_1,\nu_2)}\right)\right]. \qquad (12) \end{aligned}$$

The certainty equivalent is the smallest amount of money the investor is willing to receive with certainty at the horizon in exchange for the possibility to invest in the stock market. In formula, the certainty equivalent for a given level of investment costs, $CEQ_{(\nu)}$, is defined by the relation

$$u\left(CEQ_{(\nu)}\right) = \sup_{\pi} E\left[u\left(X^{(\nu,\pi)}(T)\right)\right].$$

Since by (2) the wealth process is linear in initial wealth we obtain

$$u\left(CEQ_{(\nu_2)}\right) = u\left(\left(1 + ICR_{(\nu_1,\nu_2)}\right) CEQ_{(\nu_1)}\right).$$

From this we conclude that for power utility the two measures are equivalent in the sense that

$$\frac{CEQ_{(\nu_2)} - CEQ_{(\nu_1)}}{CEQ_{(\nu_1)}} = \frac{\left(1 + ICR_{(\nu_1,\nu_2)}\right) CEQ_{(\nu_1)} - CEQ_{(\nu_1)}}{CEQ_{(\nu_1)}} = ICR_{(\nu_1,\nu_2)}, \qquad (13)$$

i.e. the relative change in certainty equivalents equals the indifferent compensation ratio. This is another appealing feature of power utility. The indifferent compensation ratio given by (12) can easily be calculated by (13) and (11). We get

$$ICR_{(\nu_1,\nu_2)} = \exp\left\{\frac{1}{2}\frac{T}{\gamma}\frac{\nu_2^2 - \nu_1^2 + 2(\mu - r)(\nu_1 - \nu_2)}{\sigma^2}\right\} - 1. \qquad (14)$$

Note that ICR is increasing in excess return, $\mu - r$, and decreasing in volatility, σ. Thus the more attractive stocks are (higher return / lower volatility) the higher should the compensation be. This is perhaps somewhat surprising as one might intuitively think that the impact of high investment costs would be smaller if the expected return were higher, and consequently that one would be willing to accept a smaller compensation to move from a low level of costs to a higher level of costs. In fact, the conclusion is just the opposite! This is due to the fact that for the investor we are considering a change in asset characteristics also changes the optimal allocation. Hence we are comparing the effect of costs on different allocations which hampers intuition.

Since, by assumption the investment costs are positive and smaller than the excess return offered by stocks, $0 \leq \nu_1, \nu_2 \leq \mu - r$, it follows from (11) that ICR is positive if and only if ν_2 is smaller than ν_1. Hence, one should be compensated to be willing to accept higher costs. In this respect the ICR measure conforms with intuition.

We are now in a position to quantify the impact of investment costs using the indifferent compensation measure given by (12). Consider the base case example of Subsection 3.1 with one unit initially invested and investment costs $\nu_1 = 1.4\%$ and $\nu_2 = 0.6\%$, respectively. We get

$$ICR_{(\nu_1,\nu_2)} = \frac{CEQ_{(\nu_2)} - CEQ_{(\nu_1)}}{CEQ_{(\nu_1)}} = \frac{5.66 - 4.54}{4.54} = 0.248, \tag{15}$$

i.e. the investor profits a gain equal to 24.8 percent of the initial wealth when investment costs decrease from ν_1 to ν_2. Tables 1–3 show the impact of investment costs in terms of ICR for varying parameter values. The range of risk aversions used in Table 2 corresponds to an optimal equity allocation of, respectively, 20%, 40%, 60%, 80% and 100% in the base case example with investment costs of $\nu_1 = 1.4\%$.

As illustrated in Table 1 the compensation varies greatly with the underlying market characteristics. In the most extreme case ($\sigma = 14\%$, $\mu = 11\%$) the investor demands a compensation of almost two dollars for each dollar invested to be willing to accept higher costs. In this case stocks are very attractive and the optimal allocations are highly leveraged ($\pi_1^* = 311\%$, $\pi_2^* = 349\%$), which amplifies the effect of the cost difference. For more realistic levels of capital market parameters the required compensation is still substantial, but more in line with the base case example.

Investors with different risk aversions require different compensation. Comparing the columns of Table 2 we see that the required compensation is approximately linear in the corresponding stock allocation; the allocation to stocks is five times higher in the rightmost column compared to the leftmost column. We also see from Table 2 that the compensation is approximately linear in the time horizon.

Finally, Table 3 illustrates the effect of different levels of cost. Note that in the upper right corner of the table the "low" costs are in fact higher than the "high" costs, and hence the compensation is negative. We see that the effect of a given difference in costs is highest when the absolute level is low. For example, the investor under consideration requires a compensation of 30.5% to accept costs of 0.8% rather than no costs, while he requires a compensation of "only" 19.4% to accept costs of 2.0% rather than costs of 1.2%.

4.1 Financial value of investment costs

In Subsection 3.2 we considered the change in rate of return induced by a change in investment costs and compared the naive change given by (5) with the change the VaR-investor experienced given by (7). For the utility optimizing investor it seems natural to compare the indifferent compensation ratio with the financial value of expected accumulated investment costs over the horizon. The financial value of investment costs over the horizon is defined as

$$F(\nu, \pi) \equiv E^Q \left[\int_0^T e^{-rt} X(t) \pi \nu dt \right],$$

Volatility	Expected stock return (μ)				
(σ)	4.5%	5%	7%	9%	11%
14%	7.8%	16.3%	57.2%	112.4%	187.2%
17%	5.2%	10.8%	35.9%	66.7%	104.5%
20%	3.8%	7.7%	24.8%	44.7%	67.7%
23%	2.8%	5.7%	18.2%	32.2%	47.8%
26%	2.2%	4.5%	14.0%	24.4%	35.8%

Table 1: Indifferent compensation ratio (ICR) for varying capital market parameter values. Other parameters are kept at their base case values, i.e. $r = 3\%$, $T = 40$, $\gamma = 1.833$, $\nu_1 = 1.4\%$ and $\nu_2 = 0.6\%$.

Horizon	Risk aversion (γ)				
(T)	3.25	1.625	1.0833	0.8125	0.65
1	0.2%	0.4%	0.6%	0.7%	0.9%
5	0.9%	1.9%	2.8%	3.8%	4.7%
10	1.9%	3.8%	5.7%	7.7%	9.7%
20	3.8%	7.7%	11.7%	15.9%	20.3%
40	7.7%	15.9%	24.8%	34.4%	44.7%

Table 2: Indifferent compensation ratio (ICR) for varying horizon and risk aversion. Other parameters are kept at their base case values, i.e. $r = 3\%$, $\mu = 7\%$, $\sigma = 20\%$, $\nu_1 = 1.4\%$ and $\nu_2 = 0.6\%$.

where E^Q denotes expectation with respect to the so-called risk-neutral, or pricing, measure Q. Using that the expected return of stocks equals r under Q, we get from (2)

$$E^Q\left[X(t)\right] = x_0 \exp((r - \pi\nu)t),$$

and thereby

$$F(\nu, \pi) = x_0 \left(1 - \exp(-\pi\nu T)\right). \tag{16}$$

This quantity represents the amount of money that an investor (with no investment costs) would need in order to replicate the cost cash flow charged by the fund. From the point of view of the fund manager this is the value of investment costs paid by the customer.

Next, we define the Relative Change in Financial value of accumulated investment costs induced by lowering investment costs from ν_1 to ν_2, $RCF_{(\nu_1,\nu_2)}$, by

$$RCF_{(\nu_1,\nu_2)} \equiv \frac{F(\nu_1, \pi_1^*) - F(\nu_2, \pi_2^*)}{x_0} = \exp(-\pi_2^* \nu_2 T) - \exp(-\pi_1^* \nu_1 T). \tag{17}$$

Note that, as is the case for the indifferent compensation ratio, this value is independent of the size of the initial wealth. We interpret (17) as the direct effect and the difference between (13) and (17) as the indirect effect of investment costs.

In principle, if the customer received RCF (per unit invested) he could replicate the additional costs charged by the expensive fund, and it could be argued that he should then be indifferent to joining the cheap and the expensive fund. This argument, however, underestimates the actual

High cost	Low cost (ν_2)				
(ν_1)	0%	0.3%	0.6%	0.9%	1.2%
0.8%	30.5%	17.3%	6.3%	-2.9%	-10.5%
1.1%	42.0%	27.6%	15.6%	5.7%	-2.6%
1.4%	53.2%	37.7%	24.8%	14.1%	5.1%
1.7%	63.9%	47.4%	33.6%	22.1%	12.5%
2.0%	74.0%	56.4%	41.8%	29.6%	19.4%

Table 3: Indifferent compensation ratio (ICR) for varying investment costs. Other parameters are kept at their base case values, i.e. $r = 3\%$, $\mu = 7\%$, $\sigma = 20\%$, $T = 40$ and $\gamma = 1.833$.

impact of increased investment costs. First, the additional costs must be replicated *within* the fund and hence costs must be paid on the amount set aside to cover costs! In more conventional terms, the higher costs cause lost investment opportunities since the amount of money available for investment is reduced. Second, if the investor was reimbursed for the additional costs he would optimally change his investment strategy. Thus, even if the additional costs could be replicated for free he would need compensation to accept a suboptimal investment strategy in the expensive fund. The indirect effect is the sum of these two effects.

For the base case example $(\nu_1, \nu_2) = (1.4\%, 0.6\%)$ with corresponding optimal investment strategies $(\pi_1^*, \pi_2^*) = (60\%, 78.4\%)$ we get a direct effect of

$$RCF_{(\nu_1,\nu_2)} = 0.828 - 0.715 = 0.114,$$

i.e. the financial value of investment costs equals 11.4% of the initial wealth. In contrast, we got in (15) that the investor demands an increase of 24.8% of initial wealth in compensation to accept incurring the higher investment costs $\nu_1 = 1.4\%$ instead of the lower $\nu_2 = 0.6\%$. The great gap between these two values (the indirect effect) is due to lost investment opportunities and a suboptimal strategy.

In contrast to the ICR measure, RCF is not necessarily positive when the cost ratio is lowered. When the cost ratio is lowered the allocation to stocks is increased, and in some cases this increase is so large that the value of costs (F) also increase. Assuming $\nu_2 \leq \nu_1$ it is not hard to show that $RCF_{(\nu_1,\nu_2)} < 0$ when $\mu - r < \nu_1 + \nu_2$, $RCF_{(\nu_1,\nu_2)} = 0$ when $\mu - r = \nu_1 + \nu_2$, and $RCF_{(\nu_1,\nu_2)} > 0$ when $\mu - r > \nu_1 + \nu_2$. These relations can be observed in Table 4.

It is intuitively clear that the value of costs is not monotone in the cost ratio. When $\nu = 0$ the value is zero because no costs are charged, and if $\nu = \mu - r$ the value is also zero because the allocation to stocks is zero ($\pi^* = 0$). It is easy to show that the value of costs is maximized for $\nu = (\mu - r)/2$, i.e. when costs constitute half the risk premium on stocks. In the base case example this corresponds to costs of 1.5%; slightly above the rate charged by our expensive fund.

Comparing Tables 1–3 with Tables 4–6 we see that the compensation required by the investor to accept higher costs (ICR) is substantially larger than the change in financial value of costs (RCF). Typically the indirect effect (the difference) is of the same magnitude as the direct effect as measured by RCF.

In some cases we even have that RCF is negative (leftmost column of Table 4) while ICR is positive (leftmost column of Table 1). In this situation both the investor *and* the fund manager would demand a compensation to "accept" higher costs! From the point of view of the investor, the "benefit" of higher cost ratios in terms of lower financial value is outweighed by the suboptimality of the resulting strategy.

The left plot in Figure 4 compares the indifferent compensation ratio (13) with (17) for different investment costs. In other words, we compare the size of the compensation sum with

Volatility	Expected stock return (μ)				
(σ)	4.5%	5%	7%	9%	11%
14%	-7.1%	0.0%	17.7%	24.6%	25.8%
17%	-4.9%	0.0%	14.2%	22.2%	26.0%
20%	-3.6%	0.0%	11.4%	19.0%	23.8%
23%	-2.7%	0.0%	9.2%	16.0%	20.9%
26%	-2.1%	0.0%	7.5%	13.4%	18.1%

Table 4: RCF for varying capital market parameter values. Other parameters are kept at their base case values, i.e. $r = 3\%$, $T = 40$, $\gamma = 1.833$, $\nu_1 = 1.4\%$ and $\nu_2 = 0.6\%$.

Horizon	Risk aversion (γ)				
(T)	3.25	1.625	1.0833	0.8125	0.65
1	0.1%	0.2%	0.4%	0.5%	0.6%
5	0.6%	1.2%	1.8%	2.4%	2.9%
10	1.2%	2.4%	3.5%	4.5%	5.5%
20	2.4%	4.5%	6.5%	8.3%	9.9%
40	4.5%	8.3%	11.4%	13.9%	15.9%

Table 5: RCF for varying horizon and risk aversion. Other parameters are kept at their base case values, i.e. $r = 3\%$, $\mu = 7\%$, $\sigma = 20\%$, $\nu_1 = 1.4\%$ and $\nu_2 = 0.6\%$.

the relative change in the financial value of accumulated costs over the horizon, relative to the size of the initial wealth. The fact that for reasonable values of investment costs the effect of higher costs seems to (more than) double is quite surprising. As mentioned, one could refer to the phenomenon as *The double blow of investment costs*. Clearly, one wonders how much this double effect is due to the long horizon ($T = 40$) considered by the base case example. The right plot in Figure 4 illustrates that the effect is substantial even for a short term investor. In fact, for a short time investor the indirect effect is half the size of the direct effect.

4.2 The importance of investment costs versus investment strategy

Finally we analyze the relative importance of the investment strategy and investment costs. Specifically, we consider a power utility optimizing investor who hands over his savings to a fund charging him high investment costs at a rate ν_1. In return, the fund offers a tailored investment strategy π_1^* fitted to meet the risk preferences of the client, i.e. the object given by (9) is optimized. The investor now becomes aware that another fund offers to manage his savings while only charging him investment costs at a lower rate ν_2. However, in order to offer this cheap product, the fund is organized as an investment collective meaning that all members follow the same investment strategy π_2. Obviously, if the common investment strategy π_2 exercised by the cheap fund happens to equal the tailored optimal investment strategy π_2^* it's a no-brainer — the investor should move to the cheap fund. However, how much can the collective investment strategy π_2 offered by the cheap fund differ from the optimal tailored investment strategy π_2^* for the investor still to prefer the cheap fund over the expensive fund? In formula, we look for the

High cost	Low cost (ν_2)				
(ν_1)	0%	0.3%	0.6%	0.9%	1.2%
0.8%	21.0%	11.3%	3.9%	-1.7%	-5.6%
1.1%	25.5%	15.8%	8.3%	2.8%	-1.2%
1.4%	28.5%	18.8%	11.4%	5.8%	1.9%
1.7%	30.3%	20.6%	13.1%	7.6%	3.6%
2.0%	30.9%	21.1%	13.7%	8.2%	4.2%

Table 6: RCF for varying investment costs. Other parameters are kept at their base case values, i.e. $r = 3\%$, $\mu = 7\%$, $\sigma = 20\%$, $T = 40$ and $\gamma = 1.833$.

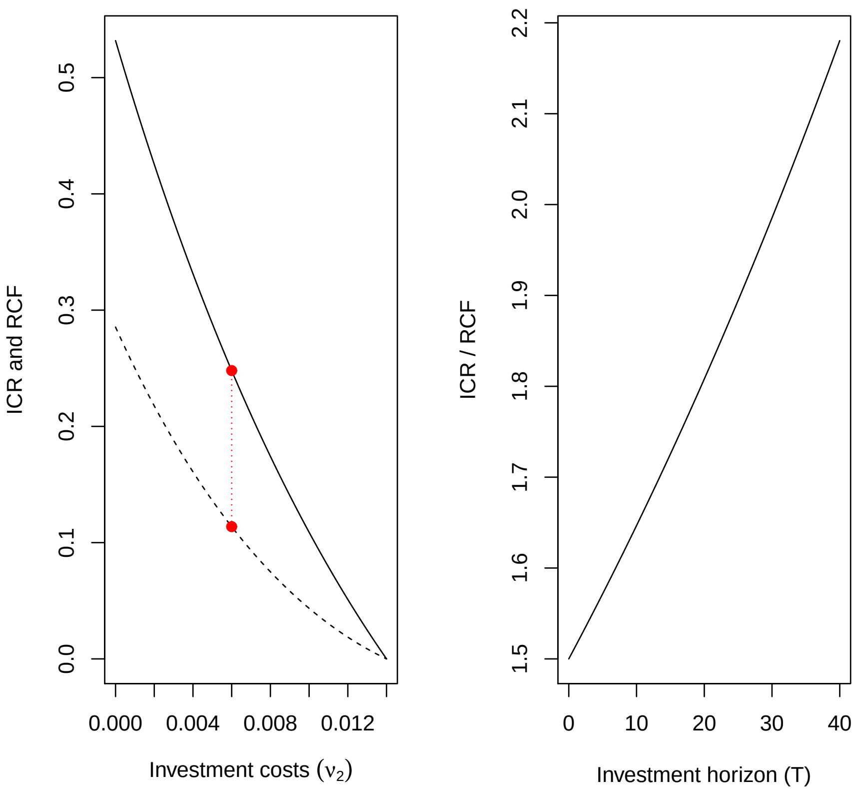

Figure 4: Illustrations of the indifferent compensation ratio (ICR) and the relative change in the financial value of accumulated investment costs (RCF). The left plot illustrates for a fixed horizon $T = 40$ the values when investment costs decreases from $\nu_1 = 1.4\%$ to $\nu_2 \in [0, \nu_1]$ (ICR solid curve, RCF dashed curve). The dots indicate the base case example where investment costs decrease to $\nu_2 = 0.6\%$. The right plot illustrates for a decrease in investment costs from $\nu_1 = 1.4\%$ to $\nu_2 = 0.6\%$ the proportion between the two measures (ICR/RCF) for varying time horizons.

investment strategies π_2 satisfying the relation

$$u\left(CEQ_{(\nu_1)}\right) = E\left[u\left(X^{(\nu_2,\pi_2)}(T)\right)\right].$$

By use of (11) we get the solution

$$\pi_2 = \frac{-b \pm \sqrt{D}}{2a}, \tag{18}$$

where $D = b^2 - 4ac$ and

$$a = \frac{\sigma}{2}(1-\gamma)\gamma,$$
$$b = -(\mu - \nu_2 - r),$$
$$c = \frac{1}{2}\frac{1-\gamma}{\gamma}\left(\frac{\mu - v_1 - r}{\sigma^2}\right)^2.$$

Note that the solution does not depend on the investment horizon.

The rightmost curve in Figure 5 illustrates, for the parameters $(\nu_1, \pi_1^*) = (1.4\%, 60\%)$ and $\nu_2 \in [0, \nu_1]$, the cost-dependent value of π_2 for which the investor is indifferent between the two funds. From left to right the next 5 curves illustrates cost-dependent values of π_2 for which the investor is 10, 20, 30, 40 and 50 percent better off being a member of the cheap fund. The line illustrates the cost-dependent optimal risk allocation π_2^*. The conclusion is surprisingly clear. If the investor is offered the lower cost rate $\nu_2 = 0.6\%$ instead of the higher cost rate $\nu_1 = 1.4\%$, he is better off almost no matter how much the investment strategy differs from his risk preferences. Any investment strategy $\pi_2 \in (0.28, 1.29)$ makes the cheap fund preferable (remember $\pi_2^* = 78.5\%$ is optimal). As seen in Figure 5 this interval shrinks when considering scenarios where the investor is 10, 20, 30, 40 and 50 percent better off. Still, the range of investment strategies is surprisingly wide. Once again, the conclusion, which is independent of the investment horizon, is very clear: The investor should be much more concerned with investment costs compared to being concerned with which investment strategy exactly meets his risk preferences.

5 Conclusion

The purpose of the paper is to quantify the total loss incurred by an investor faced with investment costs. For a CRRA utility optimizing investor faced with proportional investment costs and operating in a Black-Scholes market we find that in addition to the direct loss due to the costs themselves the investor incurs an indirect loss of similar magnitude. The indirect loss is due to lost investment opportunities and a more conservative allocation in response to (increased) costs. Thus, in a sense, the investor effectively pays costs twice. The conclusion is in line with existing results, cited in the paper, for investors with other preferences and faced with other cost structures. We are thus led to believe that this is a general result.

In contrast to existing research we quantify the impact of costs by the indifferent compensation ratio (ICR). The ratio measures the additional amount of money (per unit invested) the investor would need in compensation to be indifferent to staying in a cheap fund and moving to a more expensive fund. We derive an explicit expression for ICR, equation (14), from which a number of illuminating insights can be gained. Risk seeking investors (γ close to 0) demands a *higher* compensation than risk averse investors, and stocks with higher expected return entails *higher* compensation than stocks with lower expected return all else being equal.

We also analyze how much an investor is willing to deviate from his optimal allocation when offered lower costs; we find that he is willing to accept a very wide range of allocations indeed in exchange for lower costs. In other words, the impact of costs is far greater than the impact of the investment decision.

The analysis is carried out under the assumption that the risk aversion of the investor can be described by the power utility function (8), with known parameter γ. Perhaps, most investors do not think of their preferences towards risk in these terms. Rather, most investors have a

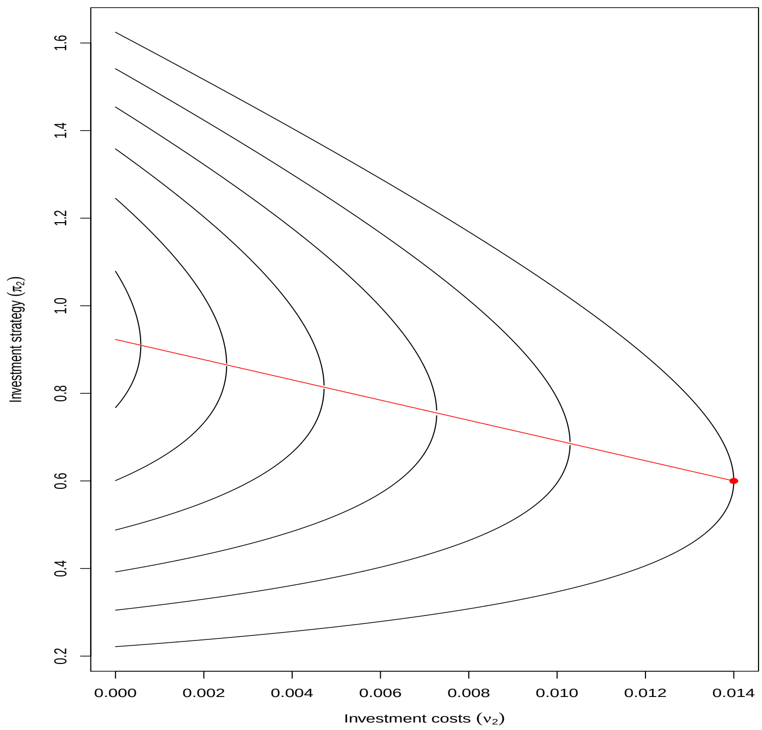

Figure 5: The curves illustrate for given investment costs $\nu_2 \in [0, \nu_1]$ the investment strategies π_2 which makes the investor (from right to left) 0, 10, 20, 30, 40 and 50 percent better off compared to being charged $\nu_1 = 1.4\%$ while using the optimal investment strategy $\pi_1^* = 60\%$. The line illustrates the cost-dependent optimal stock allocation.

preference for a certain stock allocation in a given market. Given a certain preferred stock allocation, the risk aversion parameter γ can be inferred, and we can analyze how the investor should rationally act under changed market assumptions, and particularly under changed investment costs.

We assume throughout that the investor has knowledge of the true market parameters, i.e. the rate of return and volatility. However, in real-life the investor would need to estimate these quantities from market data, and consequently he would in general be pursuing a suboptimal strategy. It can be argued that this uncertainty should be included in the analysis. On the other hand, since the impact of costs is so much bigger than the impact of the investment decision we believe that the conclusions would be essentially the same even if investors were generally following suboptimal strategies.

The analysis of the paper is carried out under the assumption that the pre-cost expected stock return is the same for all funds. There is overwhelming academic evidence for this assumption (see references quoted in the Introduction), but it is of course still debatable. However, the analysis also applies to the situation where we assume that funds can indeed generate (some) excess return. The only modification is that the cost should then be interpreted as the net-cost, i.e. the cost in excess of the assumed excess return. The main conclusion still stands; the indirect effect of costs is of the same magnitude as the costs themselves.

References

S. Basak and A. Shapiro. Value-at-risk-based risk management optimal policies and asset prices. *The Review of Financial Studies*, pages 371–405, 2001.

M. M. Carhart. On persistence in mutual fund performance. *Journal of Finance 52*, pages 57–82, 1997.

J. C. Cox and C.-F. Huang. Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory 49*, pages 33–83, 1989.

E. F. Fama and K. R. French. Luck versus skill in the cross-section of mutual fund returns. *Journal of Finance 5*, pages 1915–1947, 2010.

J. Gil-Bazo and P. Ruiz-Verdú. The relation between price and performance in the mutual fund industry. *Journal of Finance 64(5)*, pages 2153–2183, 2009.

M. Guillén, S. F. Jarner, J. P. Nielsen, and A. M. Pérez-Marín. Risk-adjusted impact of administrative costs on the distribution of terminal wealth for long-term investment. *The Scientific World Journal, http://www.hindawi.com/journals/tswj/2014/521074/*, 2014.

J. Huang, C. Sialm, and H. Zhang. Risk shifting and mutual fund performance. *Review of Financial Studies 24(8)*, pages 2575–2616, 2013.

K. Janeček and M. Sîrbu. Optimal investment with high-watermark performance fee. *SIAM Journal on Control and Optimization 50(2)*, pages 790–819, 2012.

I. Karatzas, J. P. Lehoczky, and S. E. Shreve. Optimal portfolio and consumption decisions for a 'small investor' on a finite horizon. *SIAM Journal on Control and Optimization 27*, pages 1157–1186, 1987.

A. Khorana, H. Seraes, and P. Tufano. Mutual fund fees around the world. *Review of Financial Studies 22(3)*, pages 1279–1310, 2008.

D. K. Malhotra and R. Mcleod. An empirical analysis of mutual fund expenses. *The Journal of Financial Research 20*, pages 175–190, 1997.

B. G. Malkiel. Returns from investing in equity mutual funds 1971 to 1991. *The Journal of Finance 50*, pages 549–572, 1995.

R. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics* **51**, pages 247–257, 1969.

R. Merton. Optimum consumption and portfolio rules in a continuous-time model case. *Journal of Economic Theory 3*, pages 373–413, 1971.

J. Palczewski, R. Poulsen, K. R. Schenck-Hoppé, and H. Wang. Dynamic portfolio optimization with transaction costs and state-dependent drift. *European Journal of Operational Research* doi:10.1016/j.ejor.2014.12.040, 2013.

L. S. G. Rogers. *Optimal Investment*. Springer, Berlin Heidelberg, 2013.