REPLICATING SUCCESS

ADOPTING THE VALUE-ADDED METHOD TO ASSESS MANAGERIAL SKILL IN LUXEMBOURG AND THE REST OF EUROPE

Didrik Kristiansson

Vincent Holstensson

Bachelor Thesis

Stockholm School of Economics

2024

Replicating Success: Adopting the Value-Added Method to Assess Managerial Skill in Luxembourg and the rest of Europe.

Abstract:

This study examines the managerial skill for Luxembourg's equity mutual fund market by using Berk and van Binsbergen's (2015) value-added method and using the rest of Europe as a comparison. In contrast to prior literature the results show that the managers in Luxembourg and the rest of Europe fail to persistently add value and to some extent even destroy it. Additionally, when investigating the investors' ability to predict skill, through managerial compensation, the results indicate that they are incapable of predicting performance. These results refute our hypothesis that Luxembourg fund managers could leverage from the benefit of a financial hub. Thus, the study concludes the inexistence of managerial skill in Luxembourg and the rest of Europe.

Keywords:

Mutual funds, Managerial skill, Luxembourg, Europe, Value-added

Authors:

Didrik Kristiansson (25462)

Vincent Holstensson (25313)

Tutor:

Ran Xing, Assistant Professor, Stockholm Business School.

Examiner:

Adrien D'Avernas, Assistant Professor, Department of Finance.

Bachelor Thesis
Bachelor Program in Business and Economics
Stockholm School of Economics
© Didrik Kristiansson and Vincent Holstensson, 2024

1. Introduction

The mutual fund market of Luxembourg has grown to become one of the largest investment fund centres in Europe. It is well known that Luxembourg offers beneficial tax exemptions for funds, and has a legal and regulatory environment that attracts fund managers. Given the large growth, the question arises whether managers in Luxembourg create value. Several reasons exist to support such a question. Prior studies using American data have concluded that investors invest more capital in funds that previously have outperformed their peers (Berk and van Binsbergen, 2015). Thus, as there is a large capital flow to Luxembourg managers, and if investors anticipate to compensate high-performing managers, then it follows that the Luxembourg fund managers, on average, ought to create positive value.

Additionally, Christoffersen, Susan, and Sarkissian's (2009) study investigates managerial performance in financial hubs and finds that fund managers in financial centres outperform other managers. Since Luxembourg has become a financial centre that attracts investor capital and fund managers it can be further expected that they should create value. Therefore, we hypothesise that Luxembourg fund managers will leverage the implications of financial hubs and add value.

To test our hypothesis we chose the value-added approach of Berk and van Binsbergen (2015). The value-added is defined as a fund's gross excess return over a chosen excess benchmark return multiplied by assets under management (AUM). This method has had a significant impact on the discussion regarding managerial skill by challenging the usage of alpha measurements alone to determine managerial performance. The argument stems from the fact that a manager yielding a higher gross alpha is not necessarily more skilled than another manager as fund size is not taken into account. Moreover, we do not only evaluate managers using the traditional risk factor model, Fama-French-Carhart (FFC) factor correction for risk, put forward by Fama and French (1995) and Carhart (1997). This is because the model is difficult to trade and might not have been marketed at the time. Thus, we also evaluate managers by comparing their performance to a benchmark set of index funds. For this study, we carefully reviewed Berk and van Binsbergen's (2015) reasoning regarding the choice of index funds and followed their approach.

Furthermore, we do not only compare the performance of Luxembourg managers' to a benchmark, but also to other mutual funds in Europe with the same investment universe. Our aim with this is to see whether there are any significant differences in performance between the fast growing mutual fund market Luxembourg and a similar area. Hereafter, we will denote Luxembourg, as the domicile of Luxembourg, and Europe, as the domicile of all European countries apart from Luxembourg.

Following the value-added approach, our results show that the average Luxembourg manager's performance results in value creation. Furthermore, the European counterparts also create value, albeit to a lesser extent. However, we have reason to believe these results are overstated. Thus, we also run a persistence test to look for the persistence of performance among top-decile managers to avoid any eventual statistical overestimation. The persistence test was enabled by sorting the managers into different deciles based on their performance during a sorting period. Based on this sorting, we then calculate the average value-added per decile for each regarded investment period. From this result, we could conclude that neither Luxembourg nor European fund managers show any persistence of performance over time. Additionally, we test if top managers outperform bottom managers and if top managers stay in the top half of deciles. Here we find that even though top managers at times outperform bottom managers for Luxembourg, they do not consistently occur in the top half. These findings challenge the support for persistence of managerial performance for both

Luxembourg and Europe. Consequently, using the same method as Berk and van Binsbergen (2015) we find deviating results, and the results contradict our hypothesis.

Given these results, we also tested whether compensation, the product of AUM multiplied with expense ratio, could be a better separator of skill than our sort of managers. Thus, when repeating the sorting process using compensation instead, the top managers showed no persistence in performance for neither Luxembourg nor Europe. However, interestingly the bottom managers did create value. This finding indicates that investors not only cannot predict skillful managers, but they undermine those who do create value.

Through this study we pursue to add additional evidence to the literature of the geographical perspective of managerial skill, by focusing on Luxembourg's managerial performance. Additionally, we strive to contribute to the literature regarding managerial skill by using the value-added approach.

2. Literature Review

Our study builds upon the contributions introduced by Berk and van Binsbergen (2015) to explore the existence of managerial skill through the value-added method. This method is derived from Berk and Green (2004). They argue, conditioned upon that skill is scarce, that net alpha is shaped by investor competition rather than the skill level of the manager. Through the value-added method, Berk and van Binsbergen (2015) find an existence of managerial skill, with top managers showing persistent performance over time.

Despite many studies being published regarding the existence of managerial skill within the mutual fund market, consensus is still to be found. Two sides of the literature have been established. On one end of the existing literature, evidence of managerial skill is found. Just as Berk and van Binsbergen (2015) find evidence, other studies using other methods and perspectives substantiate their findings. Kaniel, Lin, Pelger, and van Nieuwerburg (2023), use machine learning and find evidence that mutual fund managers can differentiate high from low-performing mutual funds, with outperformance persisting for more than 3 years. Additionally, three papers find managerial skill while emphasising the geographical perspective. The first article, by Christoffersen et al. (2009), adds key insight into this field. They find that financial centres' managers outperform their managerial counterparts, particularly among experienced managers. The second article, by Coval, Joshua, and Moskowitz (2001), finds that manager-investment proximity has a positive correlation with abnormal returns in the U.S. However, funds situated in larger cities face increased competition for information, which negatively impacts managerial performance. Lastly, Baik, Bok, Kang, and Kim (2010), also find support for a local advantage that yields higher excess returns.

On the opposing side of the literature, the contrary is found. The most prominent article representing this view is Carhart (1997). Carhart (1997) finds, by measuring net alpha from the FFC factor model, that persistence in mutual fund performance does not reflect superior stock-picking skills. Bollen and Busse (2001) also confirm the inexistence of managerial skill, by looking at the daily frequency of return data. Furthermore, Gruber (2001) confirms this view, that fund-picking skills among managers might be more due to luck than their actual abilities. Moreover, additional studies find an inexistence of managerial skill when examined at an aggregate level, yet they suggest that a minority of managers demonstrate skill. One example is by Fama and French (2010). They use a bootstrap model and find that proficient managers may be obscured in the aggregate results by the performance of managers with insufficient skills. Furthermore, Kosowski, Timmermann, Wermers, and White (2006), use a bootstrap model and find evidence that fund managers have superior stock-picking skills. However, only a sizable minority of managers demonstrate stock-picking skills that exceed their costs.

Our study aims to contribute to the existing literature in two ways. First, we aim to provide additional evidence to the geographical perspective of managerial skill, focusing on the fastest-growing mutual fund market, Luxembourg. Second, we seek to contribute to the published literature covering the usage of the value-added approach to measure the existence of managerial skill.

3. Theoretical Background

The value-added approach measures the monthly dollar value that the fund manager of the *i*th fund adds over the benchmark. To calculate the value-added we first let R_{it}^{g} denote the excess gross return of fund *i* at time *t* and R_{it}^{B} denote the excess net return of the alternative investment, *i*, an investor could have invested in at time *t*. Then the benchmark adjusted realised gross return, $R_{it}^{g} - R_{it}^{B}$, is multiplied by the real size of the fund's inflation-adjusted asset under management (AUM) at the end of the previous period, $q_{i,\,t-1}$. The value-added is also equivalent to the lagged inflation-adjusted AUM multiplied with the addition of the fee, f_{it}, and the error term, ε_{it}, for fund *i* at time *t*. The error term in this equation is also known as the net alpha. Following this, the value-added between time *t-1* and *t* is computed as

$$V_{it} \equiv q_{i,t-1} \cdot (R_{it}^{g} - R_{it}^{B}) = q_{i,t-1} \cdot f_{it} + q_{i,t-1} \cdot \varepsilon_{it} \tag{1}$$

If a manager has skill a fund should, over the course of its existence, produce a positive value-added on average. Consequently, the managerial skill is defined as

$$S_i \equiv E\left[V_{it}\right] \tag{2}$$

From this, the estimated value-added, $\widehat{S}_i$, given that the fund exists for T_i periods, follows as

$$\widehat{S}_i = \sum_{t=1}^{T_i} \frac{V_{it}}{T_i} \tag{3}$$

From this, the average value-added can be calculated in two ways depending on whether we want to account for non-surviving funds or not. If we do not account for the non-surviving funds the distribution, also called the *ex-ante* distribution, is given by

$$\overline{S} = \frac{1}{N} \sum_{i=1}^{N} \widehat{S}_i \tag{4}$$

where *N* is the number of funds. However, if we do want to account for the non-surviving funds the distribution, also called the *ex-post* distribution, is given by

$$\overline{S}_W = \frac{\sum_{i=1}^{N} T_i \widehat{S}_i}{\sum_{i=1}^{N} T_i} \tag{5}$$

Our study has two forms of the null hypothesis, a strong form and a weak form. The strong form of the null hypothesis, proposed by Fama French (2010), is that no manager possesses skill. However, if managers lack skill they will still charge fees for the service that they offer. Consequently, their compensation would only be paid by irrational investors. Accordingly, $q_{i,t} f_{i,t-1} = -E\left[q_{i,t-1}\varepsilon_{it}\right]$ implying

$$S_i = 0 \;\forall\; i \tag{6}$$

As no manager possesses skill, it follows that the average manager does not have skill either, thus:

$$\overline{S} = \frac{1}{N} \sum_{t=1}^{N} \widehat{S}_i = 0 \tag{7}$$

If all managers lack skill, differences in the cross-section of skill is not possible. Thus, under the strong form of the null hypothesis, positive value added in the past cannot predict value added in the future. We can also test Eq. (7) in isolation, which we define as the weak form of the null hypothesis. This implies that the average manager does not have skill, although it does not exclude that some individual managers may possess such skill.

The alternative hypothesis suggests that skillful managers do exist. Even if some managers have little or no skill we argue that they still can invest in indexes as a minimal effort, thus not pursuing value destruction. Consequently, from this way of reasoning the skill, S_i, would be greater than or equal to zero,

$$S_i \geq 0 \; \forall \; i \tag{8}$$

Therefore, even if the investors are irrational and potentially would accept a negative alpha the value-added still cannot be negative, as long as the managers optimise. Consequently, even an unskilled manager will not destroy value. Then it follows that the average manager must possess skill since the expected value-added will never be less than zero,

$$\overline{S} > 0 \tag{9}$$

4. Choice of benchmarks

To measure the skill of fund managers an alternative investment opportunity is required. We call the alternative investment opportunity the benchmark of which we use two, a set of index funds and the FFC factor correction for risk. A set of index funds as a benchmark is chosen in accordance with recent studies as prior literature has debated whether risk factors accurately can correct for risk. The FFC factor correction for risk is used for comparative reasons, it has been widely used in previous studies and could still prove as a valid risk model.

For the index benchmark, there are several conditions that an index fund should fulfil in order to be considered. First, the index fund must be tradeable and of low cost. It cannot be a bond, real estate, or index fund that was spanned by existing funds. Optimally the index fund should also outperform the other available similar indexes over time and have the same investment area as the mutual fund.

For this study, we have chosen a global investment area. Our reason is to not constrain managers' ability to express their skill by limiting the options available to them. However, for comparative reasons, we must restrict it in some way. It would be unwise to compare Luxembourg with Europe if their investment area were vastly different. Thus we only include funds that invest globally. For this reason, we have decided to use the same index benchmark as Berk and van Binsbergen (2015). They came to the conclusion that a set of Vanguard index funds provide a reasonable description of what investors' actual alternative investment opportunity was at the time. The indexes are presented in Table 1, and the gross and net return data was gathered from the *Morningstar Direct* database. We acknowledge that there is a potential shortcoming with using this set of Vanguard indexes in our study. These sets of indexes were not introduced in Europe until 1998 and only became widely available around 2009. However, as is further discussed in section 5.1, the vast majority of mutual funds in our data set are born after 1998.

Table 1: Benchmark Index funds

The table shows the list of the 11 indexes that is used to calculate the index benchmark. The table also shows its respective ticker and the inception dates gathered from the *Morningstar Direct* database.

Index	Ticker	Inception Date
S&P 500 Index	VFINX	08/31-1976
Extended Market Index	VEXMX	12/21-1987
Small-Cap Index	NAESX	10/30-1960
European Stock Index	VEURX	06/18-1990
Pacific Stock Index	VPACX	06/18-1990
Value Index	VVIAX	11/13-2000
Balanced Index	VBINX	11/09-1992
Emerging Markets Stock Index	VEIEX	05/04-1994
Mid-Cap Index	VIMSX	05/21-1998
Small-Cap Growth Index	VISGX	05/21-1998
Small Cap Value	VISVX	05/21-1998

These indexes are then used to calculate the benchmark return, R^B_{it}. For this, we let R^j_t denote the excess return earned by investors in the *j*th benchmark index fund at time *t*, and the

number of funds offered in the index benchmark at time t as $n(t)$. Consequently, the index benchmark return is given by

$$R_{it}^{B} = \sum_{j=1}^{n(t)} \beta_i^j R_t^j \tag{10}$$

where β_i^j is extracted from the linear projection of the active mutual fund onto the set of benchmark index funds. As can be noted from Table 1, the inception date varies between the vanguard indexes. Thus, we sort the indexes based on how long they have been in existence and then create an orthogonal basis to make them uncorrelated. The orthogonal basis is constructed by projecting the *n*th index fund onto the orthogonal basis of the preceding funds, *n-1*, for the months fund *n* exists. For the time periods when fund *n* does not exist, zeros are inserted. We also construct an augmented basis where the zeros in the orthogonal basis are replaced with the mean value of the orthogonal basis fund when it does exist. This method follows from Berk and van Binsbergen (2015) where beta coefficients, β_i^j in Eq. (10), are calculated using the augmented basis and the benchmark return using Eq. (10) and the orthogonal basis.

The other benchmark that we use is the FFC factor model to adjust for risk. The FFC factor model includes 4 factors, market risk (MKT), size (SMB), value (HML), and one-year anomaly return (UMD), downloaded from *Wharton Research Data Service*. This has been the standard way to measure the riskiness of a manager's portfolio. In this case, for fund *i* the benchmark return is a portfolio of equivalent risk using FFC factor portfolios

$$R_{it}^{B} = \beta_i^{mkt} MKT_t + \beta_i^{smb} SMB_t + \beta_i^{hml} HML_t + \beta_i^{umd} UMD_t \tag{11}$$

This projection includes the four risk factors and gives individual risk exposures, β_i, of the *i'*th mutual fund, similar to that of the linear projection in Eq. (10).

As FFC factor adjustment for risk has been the standard, we feel a quantitative validation of the usage of the chosen Vanguard index is in order. Thus, we regress the FFC factor portfolios on the orthogonal basis set of our indexes. To make comparisons fair we have ensured that the same interest rate is used for all operations throughout our study. The results can be seen in Table 2. The table shows us a similar pattern as reported by Berk and van Binsbergen (2015) where all factors possess positive alphas. However, there are some slight deviations, specifically the MKT factor is statistically significant while the HML factor is not. Nevertheless, despite these deviations the economic significance is negligible. Additionally, the subsequent R-squared value is high for all factors except for the UMD factor, which is not surprising since a momentum strategy requires active trading not feasible for index funds. In summary, the FFC mimicking portfolios have been better investment opportunities than the indexes offered during the study period, especially the UMD factor. However, as these factors are not tradable they cannot be interpreted as a real benchmark. Still, there is reason to include them for the arguments provided earlier.

Table 2: Net alpha of FFC portfolios regressed on indexes

The table shows the extracted alpha in basis points, t-statistic of the alpha, and adjusted R-squared for each of the Fama French Carhart factors regressed on the 11 indexes presented in Table 1.

	MKT	SMB	HML	UMD
Alpha (bp/month)	1.82	9.67	12.17	60.20
t-statistic	1.96	2.22	1.25	2.89
Adjusted R^2	99.8%	92.1%	65.2%	23.4%

5. Data set and description

5.1 Data set

The European survivorship bias-free equity mutual fund data we used was extracted from the *Morningstar Direct* database from January 1990 to December 2023. Our choice of this time frame is attributed to when the mutual fund market in Luxembourg began its dominance. The retrieved mutual fund data contained funds domiciled in both Luxembourg and Europe, amounting to a total of 110,240 funds. Although a widely used data set, some corrections of the shortcomings in the data have been required for our analysis, these are presented below. Further details about the interest rate and the extracted variables are explained in Appendix B.

First, our data set includes index funds and these have to be excluded to assess managerial skill. We did this in two steps. Morningstar has its classification of whether a fund is an index fund or not, but we found inconsistencies in this classification. More specifically, some funds not classified as index funds were actually index funds. Thus, we investigated funds identified as index funds and generated a list of common words typical for index funds, see Table 1A in Appendix A. We then removed all funds including these words. After this process, we are left with 707 funds identified as index funds, which we then drop. We realise this process might not capture all index funds and risk falsely removing some mutual funds. However, it ensures that more index funds are excluded than if only Morningstar's classification were to be used. Moreover, we dropped all funds that do not invest in a global investment area and split our data into two sets, one containing Luxembourg-domiciled funds and the other Europe-domiciled funds.

For the empirical analysis, we require an expense ratio for each fund at time t. This is not directly reported by Morningstar which is a problem. To bypass this issue we utilise the fact that both gross return and net return are reported. Thus, to compute the expense ratio for fund i at time t we subtracted the net return with the gross return at time t. However, in this process, we found some cases where the expense ratio became negative. In those cases, we replaced the value with the average non-negative expense ratios up until that time for that fund. Correspondingly, the gross return data was adjusted in these cases. Similarly, we found gaps in AUM for some funds in our data set. In these cases, we fill the missing observations with the most recent observation in the past.

Another important correction of the data set is to take subclasses into account. Morningstar's database reports subclasses of a fund as separate where the difference mainly is in the fees charged. Thus if we were to not combine these the statistical significance would artificially increase for any effect we find. Fortunately, Morningstar reports *FundIds* for each fund and thus combining subclasses is an easy process. Subsequently, we aggregate subclasses with the same *FundId* into a single fund.

Next, we remove funds containing no data for at least one variable of interest to our analysis and inflation adjust each AUM observation for each fund to January 1, 2000 dollars. The inflation adjustment is described in further detail in section 10.2 in Appendix B.

Unfortunately, for most funds in our data set AUM and gross return begin to be reported later than net return. We decided to drop fund observations where the fund has less than $5 million after the inflation adjustment to the year 2000 dollars. As AUM starts to be reported later than net return some data is lost, mostly before the mid 2000s. However, if managers are skilled, they should not only be able to express skill in the early years of the fund but throughout its whole lifetime. Thus, this should not have any significant effect on our results. Finally, we drop any fund with less than 24 months of consecutive data. These

steps in the data project reduced the total number of funds for the Luxembourg data set to 2233 and the European data to 4330.

5.2 Data description

Since the extracted data sets comprise multiple countries, it is of interest to determine what domiciles are most prevalent. Table 3 presents descriptive statistics for Luxembourg and Europe as well as the top 5 most frequently occurring domiciles in the European data. The top 5 countries in Europe are The United Kingdom, Ireland, Germany, France, and Spain, while the category "Other" denotes the remaining countries. From this table, we see that Luxembourg constitutes the relative majority of the total number of funds, 34.02%, compared to other domiciles. Also, no clear differences exist between domiciles in the summary statistics.

Table 3: Descriptive statistics per domicile

The table shows the excess gross return for Luxembourg, The United Kingdom, Ireland, Germany, France, Spain, and the remaining countries aggregated to "Other". Across funds, the table shows the mean, median, standard deviation from the mean, highest, and lowest average excess return in each domicile. As well as, the percentage of total observation in each domicile and the percentage of Europe for the European domiciles excluding Luxembourg.

			Top European domiciles					
Country	Luxembourg	Europe	The U.K.	Ireland	Germany	France	Spain	Other
Mean	0.30%	0.23%	0.34%	0.30%	0.13%	0.18%	0.08%	0.27%
Std. dev.	0.44%	0.46%	0.35%	0.49%	0.49%	0.41%	0.42%	0.48%
Highest	2.13%	2.94%	2.26%	2.93%	2.10%	1.71%	2.64%	1.88%
Median	0.38%	0.29%	0.36%	0.37%	0.20%	0.23%	0.09%	0.34%
Lowest	-3.39%	-5.34%	-1.47%	-3.33%	-3.05%	-1.83%	-2.37%	-5.34%
Percentage of total	34.02%	65.98%	9.43%	9.23%	7.77%	7.76%	7.60%	24.18%
Percentage of Europe	-	-	14.30%	14.00%	11.78%	11.76%	11.52%	36.65%
Number of funds	2233	4330	619	606	509	510	499	1587

Before presenting our results, a concern is that the average value-added, S_i, can be affected if the growth in the industry's AUM is constituted from increases in fund size rather than the number of funds. This concern is alleviated by Figure 1 and Figure 2, which demonstrate the inflation-adjusted AUM, for the displayed percentiles over time. The solid black line depicts the logarithm of the total number of funds. We note that the median inflation-adjusted AUM has been stable throughout the time period. Consequently, we can conclude that the growth in the industry's AUM is driven by the number of funds and not by increases in fund size for both Luxembourg and Europe. Another concern is that the distribution of available data between Europe and Luxembourg is different. Such concerns are relieved in Figure 1A and Figure 2A in Appendix A as both data sets follow a similar pattern over time and the percentage of data availability is nearly identical.

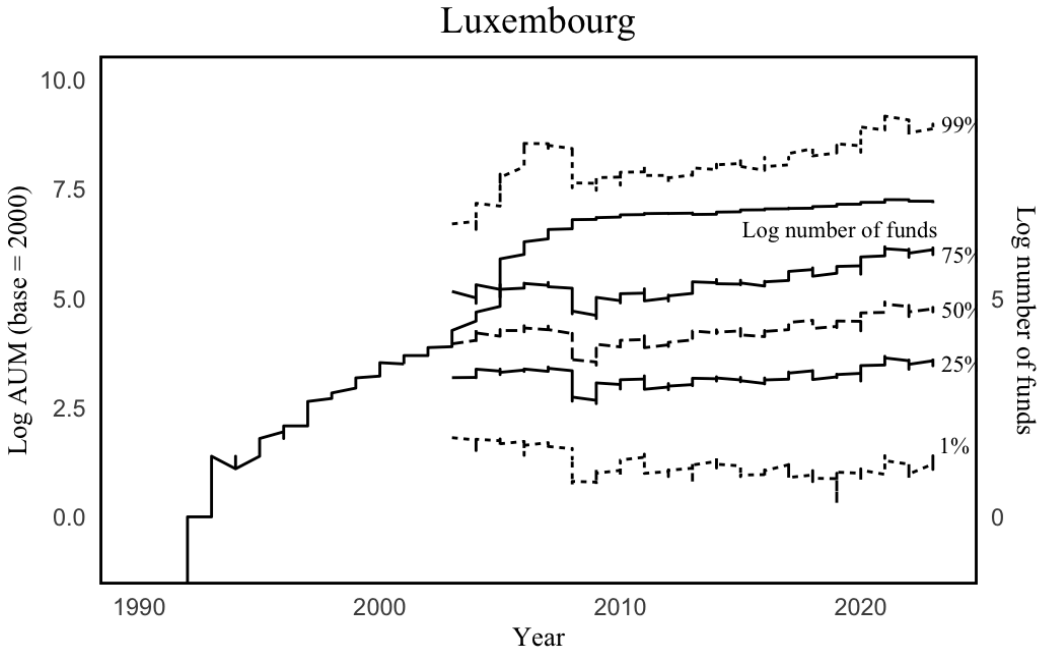

Fig. 1. Distribution of fund size for Luxembourg: The graph displays the distribution of the logarithm AUM in millions of dollars (year 2000 dollars) of each percentile (1st, 25th, 50th, 75th, 99th) at each point in time for the Luxembourg data set. The black solid line depicts the logarithm of the number of funds for the data set.

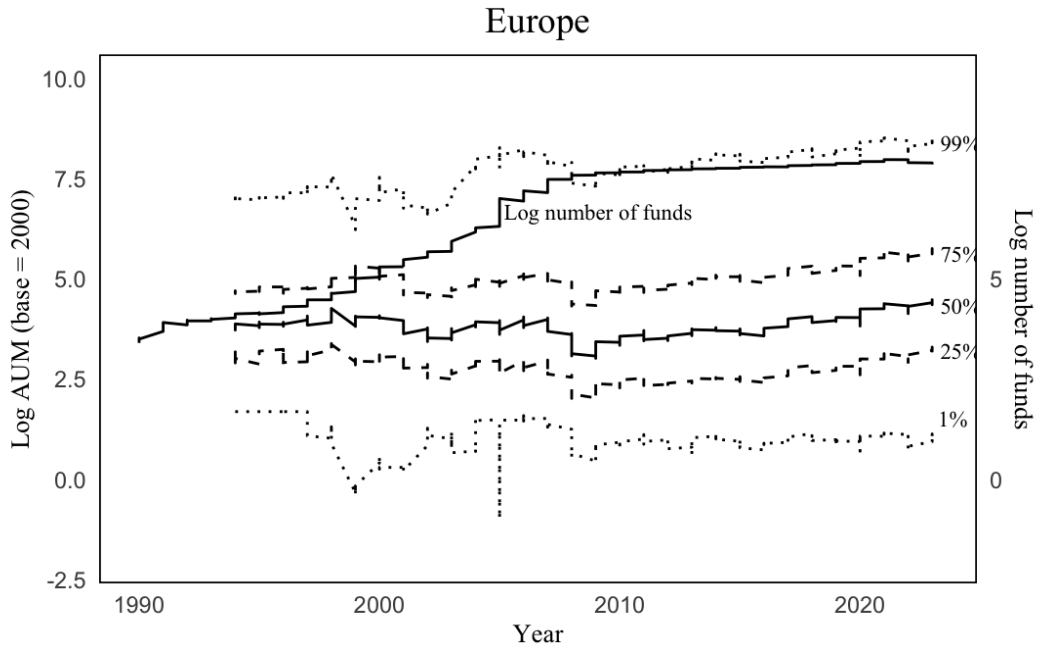

Fig. 2. Distribution of fund size for Europe: The graph displays the distribution of the logarithm AUM in millions of dollars (year 2000 dollars) of each percentile (1st, 25th, 50th, 75th, 99th) at each point in time for the European data set. The black solid line depicts the logarithm of the number of funds for the data set.

6. Results

6.1 Measuring skill and performance persistence

To test our hypothesis we calculate the cross-sectional mean and cross-sectional weighted mean of s_i from Eq. (4) and Eq. (5) for the Luxembourg and European managers. Table 4 shows the subsequent results for our sample data. For Luxembourg-domiciled funds, the cross-sectional weighted mean is positive and statistically significant. This indicates that the average Luxembourg mutual fund manager adds $350,000 monthly or $4.2 million annually (in Y2000 dollars). Furthermore, the *ex-ante* distribution, Eq. (4), shows a $160,000 monthly value-added. This is a lower estimate than the *ex-post*, which is reasonable given that unskilled managers exit sooner. Comparatively, the European fund managers add, on average, a statistically significant value of $260,000 per month, or $3.1 million annually. When deriving the value-added through the FFC factor specification, the conclusion for both domiciles mostly remains the same. Also, we see differences in the percentages of funds with a positive estimated value added for Luxembourg and Europe. For Luxembourg-domiciled funds, the majority of managers create value, while only 44.37% of funds in Europe had a positive estimated value added. Overall, these findings suggest that the weak form of the null hypothesis can be rejected for both Luxembourg and European managers. This indicates that, on average both managers exhibit skill, but arguably Luxembourg creates more value.

Table 4: Cross-sectional and cross-sectional weighted mean for Luxembourg and Europe
The table shows the cross-sectional mean, $\bar{s}$, from Eq. (4) and the cross-sectional weighted mean, $\bar{s}_w$, from Eq. (5) for every fund in Luxembourg and the European data set. The standard error of the mean and t-statistics along with the percentage under zero for the cross-sectional mean and cross-sectional weighted mean are presented. The mean is calculated for both the FFC factor specification and the Index benchmark. The results are presented in millions of Y2000 dollars/month.

Category	Index Benchmark		FFC	
	Luxembourg	Europe	Luxembourg	Europe
Cross-sectional weighted mean	0.35	0.26	0.20	0.15
Standard error of the weighted mean	0.03	0.03	0.04	0.04
t-statistic (Weighted)	10.07	9.46	5.61	4.28
Cross-sectional mean	0.16	0.16	0.02	0.06
Standard error of the mean	0.03	0.03	0.04	0.04
t-statistic	4.52	5.77	0.43	1.82
Percentage under zero	45.16%	55.63%	50.16%	64.83%
Total number of funds	2232	4305	2233	4313

However, before making conclusions prudence should be exercised. Similar to Berk and van Binsbergen (2015), there is a risk that our t-statistics are overstated. The overestimation stems from two reasons. Firstly, it is likely that a correlation in value-added exists between funds. Secondly, the value-added distribution features excess kurtosis.

Given these econometric shortcomings of our t-statistics, we utilise the assumption from the strong null hypothesis to derive an alternative measure of statistical significance.

Under the null, there cannot be any persistence of value-added, which indicates that top performing managers in the past cannot remain the top managers in the future. Conversely, if there is skill among managers, then relative outperformance should persist over time. Hence, a more powerful test of the strong form of the null hypothesis can be achieved by counting the number of times in the future that top managers beat bottom managers, and the number of times the top managers are in the top half. This derived distribution is Binomial $(n, \frac{1}{2})$, where n is the number of future monthly observations. As a result, an exact p-value can be calculated for each test statistic, circumventing the aforementioned problems.

To test the persistence of skill, we first sort funds into deciles based on the same inference of managerial skill as Berk and van Binsbergen (2015). Thus, to infer skill at time τ, we construct a Skill Ratio as

$$SKR_i^{\tau} \equiv \frac{\hat{S}_i^{\tau}}{\sigma(\hat{S}_i^{\tau})} \tag{12}$$

where $\hat{S}_i^{\tau} = \sum_{t=1}^{\tau} \frac{V_{it}}{\tau}$ and $\sigma(\hat{S}_i^{\tau}) = \sqrt{\sum_{t=1}^{\tau} \frac{(V_{it} - \hat{S}_i^{\tau})^2}{\tau}}$. The skill ratio at time τ is the t-statistic of the value-added estimate measured over the entire history of the fund until time τ. This time period is called the sorting period as funds are sorted into deciles based on our confidence that they will generate positive value-added during the sorting period. To get reliable sorts, each fund is required to have at least 3 years of historical data. Once funds are sorted we test their persistence for different time horizons, which we call the measurement horizons, h, by calculating the monthly value-added of each fund $\{V_{i,\tau}, \ldots, V_{i,\tau+h}\}$. In total we use 6 measurement horizons, each being a year longer than the previous. Thus, the measurement horizons span from 36 to 96 months. The amount of measurement horizons was limited due to insufficient funds for higher measurement horizons.

If we were to use the same estimate of the fund's betas for both periods then the estimation errors of funds' betas would be positively correlated between the sorting period and the measurement horizon. To avoid this bias we refrain from using information in the sorting period to estimate the betas for the measurement horizon. Thus we require sufficient amounts of data in the measurement horizon, first to reliably estimate betas and then still have enough observations left to measure performance. Therefore, we consider the shortest measurement horizon of three years, with the first m months in the measurement horizon used to estimate the betas. Because a minimum amount of months is needed to estimate the betas, we drop any fund with less than m months of gross return data. Also, to avoid selection bias we remove the first m months in the measurement horizon. Thus the measurement horizon we are left with is $\{V_{i,\tau+m}, \ldots, V_{i,\tau+h}\}$. In accordance with Berk and van Binsbergen (2015), we choose $m = 18$ for the index benchmark and $m = 6$ for the FFC factor specification. Once value-added is calculated for each fund in each decile for that horizon, funds are re-sorted according to the skill ratio based on all information until that time. This process is repeated as many times as the data allows. In the end, we are left with monthly value-added estimates for each decile, resulting in a time series.

From this time series, the above order statistics as well as the mean and standard errors for each decile across all measurement horizons are calculated. The subsequent results are less affected by the t-statistics as it is calculated using time series observations of the decile means. Thereby eliminating the effect of cross-fund correlation and substantially reducing the excess kurtosis. In Figure 3 the results for the index benchmark derived out-of-sample value-added

are shown for all measurement horizons. For Luxembourg, panel A, the results are statistically insignificant across all horizons at a 95% confidence interval for the tenth decile funds. The top decile funds are also not consistently the best-performing decile throughout the measurement horizons and at times even perform the worst. Comparatively, no decile for the European domiciled funds, in panel B, adds or destroys value. These results indicate a lack of persistence in performance for top managers in both Luxembourg and Europe when the index benchmark is used. Figure 4 plots the same results as Figure 3 but using the FFC risk factor specification instead. Not surprising, the results are similar to Figure 3. The vast majority of deciles across all measurement horizons for both Luxembourg and Europe do not create any statistically significant value added. However, one slight deviation between the results is that decile 7 for Luxembourg consistently adds value over different measurement horizons. Unfortunately, this is not an indication of persistence in skill as the sorting procedure predicted the tenth decile to employ consistency and not decile 7. Consequently, this is viewed as an anomaly and therefore we conclude that no manager shows persistence in creating value.

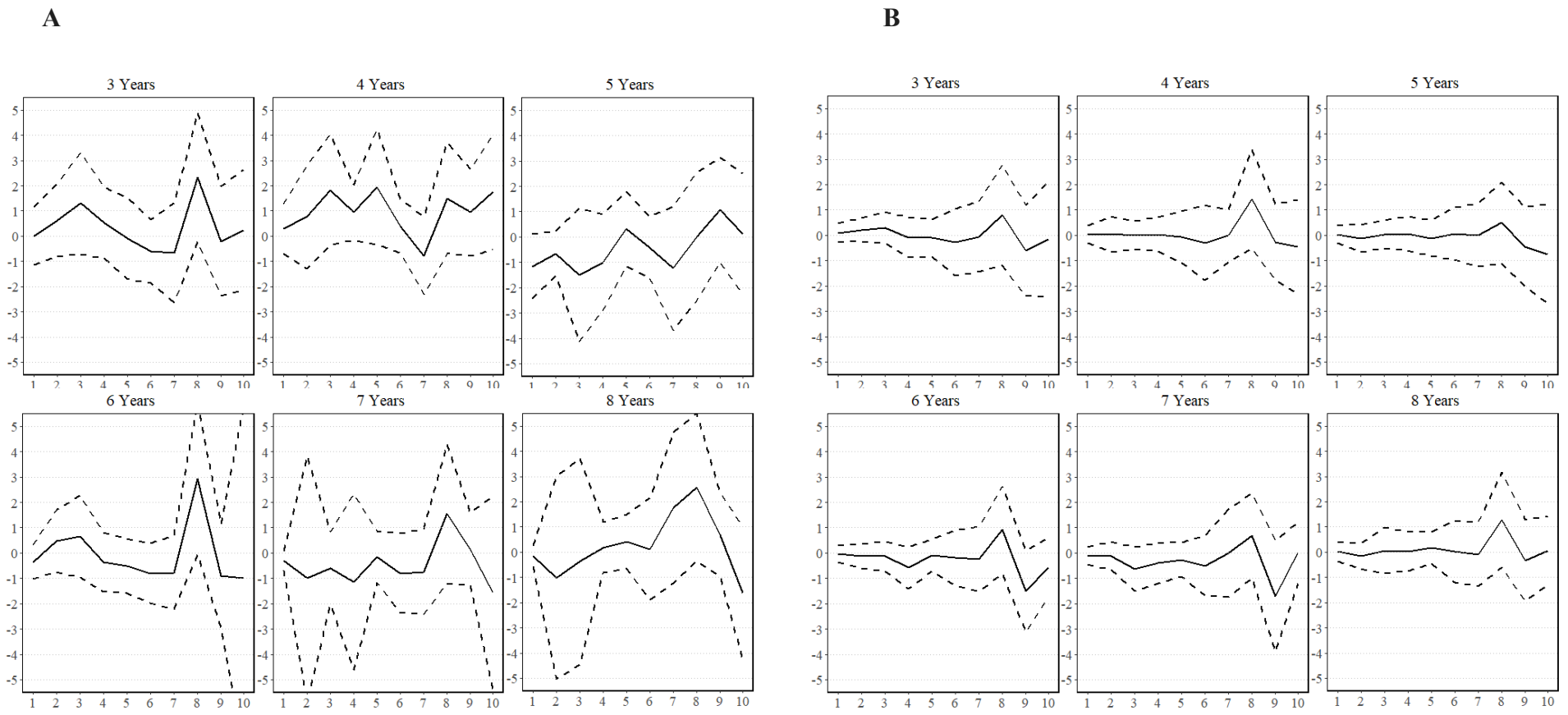

Fig. 3. Index benchmark derived out-of-sample value-added for Luxembourg and Europe: The figure displays out-of-sample value-added, $\hat{S}_i$, for each fund in millions of Y2000 dollars/month extracted using the index benchmark, sorted into deciles based on skill ratio (the solid line). Each graph shows the results for the different measurement horizons indicated by the title, and the deciles are indicated by the horizontal axes. The dotted lines indicate the two standard error bounds. Panel A shows the results for Luxembourg funds, and panel B shows the results for European funds.

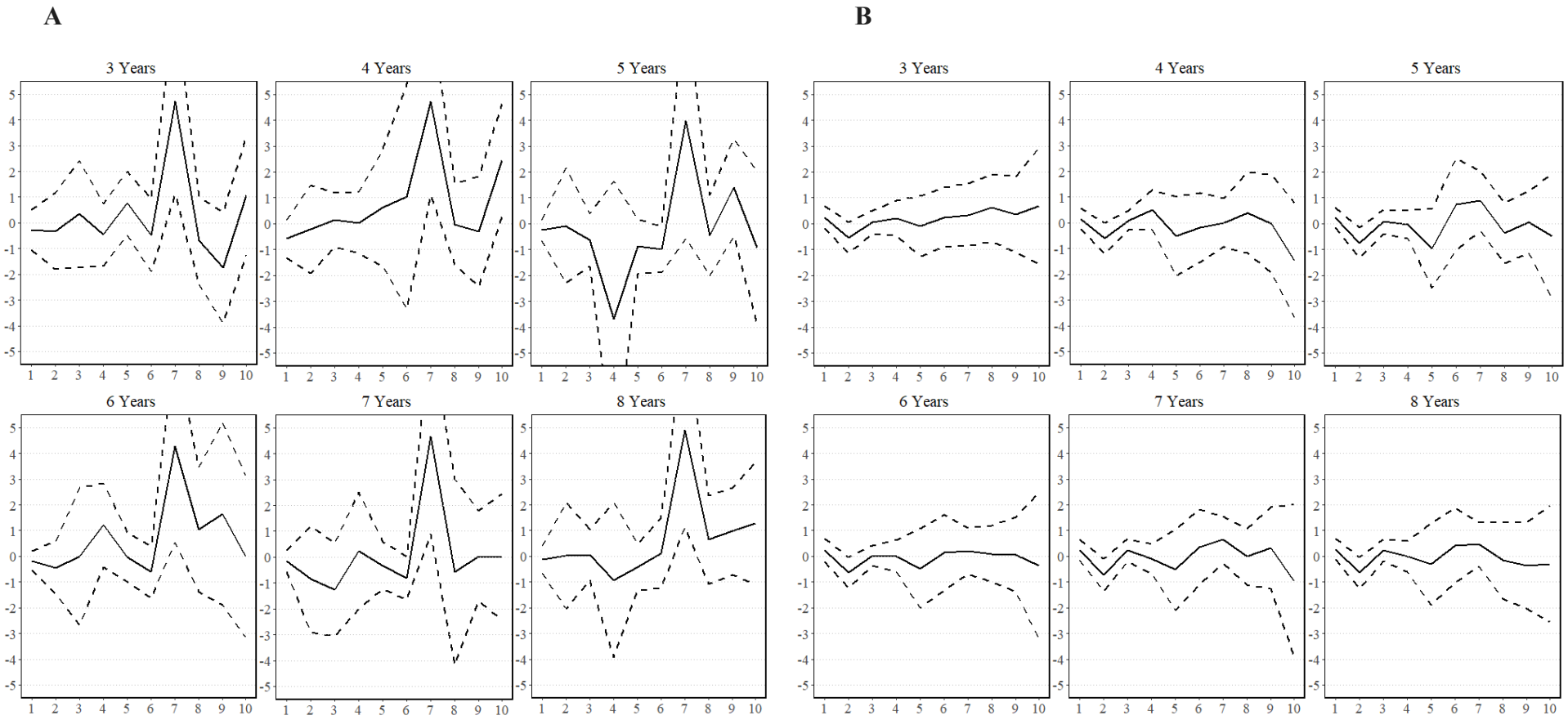

Fig. 4. FFC factor specification derived out-of-sample value-added for Luxembourg and Europe: The figure displays out-of-sample value-added, $\hat{s}_i$, for each fund in millions of Y2000 dollars/month extracted using the FFC factor specification, sorted into deciles based on skill ratio (the solid line). Each graph shows the results for the different measurement horizons indicated by the title, and the deciles are indicated by the horizontal axes. The dotted lines indicate the two standard error bounds. Panel A shows the results for Luxembourg funds, and panel B shows the results for European funds.

We next formally test the null hypothesis that the value-added of the tenth decile managers is zero or less, under the traditional asymptotic assumptions. We also measure how many times a top fund manager is in the top half and a top manager beats a bottom manager. That is, the number of times a tenth decile fund is in the top half, and the number of times a tenth decile fund beats a first decile fund. The results are presented in Table 5. In the majority of all measurement horizons for Luxembourg and Europe the top decile shows statistically insignificant value-added. However, Luxembourg-domiciled funds demonstrate consistency in top decile funds outperforming bottom decile funds as opposed to Europe. Still, Luxembourg-domiciled, with a few exceptions, and European funds fail to show consistency in staying in the top half. Thus, we cannot exclude that top managers are attributable to luck. Therefore, we conclude that both Luxembourg and European top managers show no persistence in performance and the strong form of the null hypothesis cannot be rejected.

Nevertheless, Berk and van Binsbergen (2015) conclude that the skill ratio is unlikely to be a perfect way to separate skill, but compensation, the AUM multiplied by the expense ratio, could be a better separator. Thus, we next look into the possibility that investors can determine a high-performing manager *ex-ante*.

Table 5: Performance of the top decile fund

The table shows the average value-added, derived from the FFC factor specification and Index benchmark, with its associated p-value of the top decile manager based on the skill ratio for both Luxembourg and European data sets. Additionally the frequency for top managers beating bottom and the top managers staying in the top half with their subsequent p-values are presented. All presented p-values are one-tailed.

Horizon	Value added in $ Mil		p-value (%)		Top beats bottom freq. (%)		p-value (%)		Top in top half freq. (%)		p-value (%)	
	Luxembourg	Europe	Luxembourg	Europe	Luxembourg	Europe	Luxembourg	Europe	Luxembourg	Europe	Luxembourg	Europe
Vanguard												
3	0.24	−0.16	42.18	55.59	54.32	46.80	15.35	80.02	49.19	34.15	55.85	100.00
4	1.75	−0.46	6.66	68.65	57.06	46.80	3.89	80.02	49.19	31.69	55.85	100.00
5	0.11	−0.74	46.36	77.08	55.29	52.22	9.61	28.72	48.65	39.08	61.57	99.99
6	−0.99	−0.58	61.32	83.43	58.33	45.32	1.86	89.68	52.46	31.34	27.71	100.00
7	−1.56	0.02	78.87	48.62	53.53	49.75	19.94	50.00	48.65	34.86	61.57	100.00
8	−1.61	0.07	88.41	46.25	52.94	46.80	24.50	80.02	47.03	34.51	76.89	100.00
FFC												
3	1.06	0.68	18.30	27.81	60.92	50.23	0.25	50.00	55.44	38.80	7.50	99.99
4	2.45	−1.46	1.46	90.05	64.29	43.72	0.01	96.19	59.07	33.11	0.72	100.00
5	−0.93	−0.48	73.17	65.25	57.69	50.70	2.27	44.58	53.89	37.46	15.68	100.00
6	0.00	−0.37	49.97	59.86	56.59	52.09	4.41	29.27	53.89	38.13	15.68	100.00
7	0.01	−0.95	49.68	73.50	53.85	47.91	16.76	70.73	49.74	35.12	50.00	100.00
8	1.28	−0.29	14.46	60.03	60.44	51.16	0.30	39.25	55.44	36.79	7.50	100.00

6.2 Compensation-sorted funds

We first investigate the monthly compensation per decile sorted by skill ratio, shown in Figure 5. From these graphs, we can conclude that investors reward managers with higher compensation once they have added value. This means that investors potentially are able to determine performance ex-ante. Thus, we investigate whether compensation is a better separator of fund manager performance than the skill ratio previously used. For this, we follow the same procedure to sort funds into deciles but base our sort on total compensation. That is, at the beginning of each time horizon we sort funds into deciles based on the product of current AUM and fee, and then we follow the previous procedure identically.

Figure 6 shows the results of this process for the index benchmark derived value added. The results indicate that the highest compensated managers across all measurement horizons for both Luxembourg and Europe do not create any value at a 95% confidence interval. However, some bottom decile funds create a statistically significant value-added for Luxembourg and Europe. Figure 7 follows the same pattern with first decile managers producing statistically significant values. This indicates that under-compensated managers in previous periods perform best in subsequent periods, potentially expressing skill. Thus, investors not only seem to not be able to predict skillful managers ex-ante, but instead undermine those creating value.

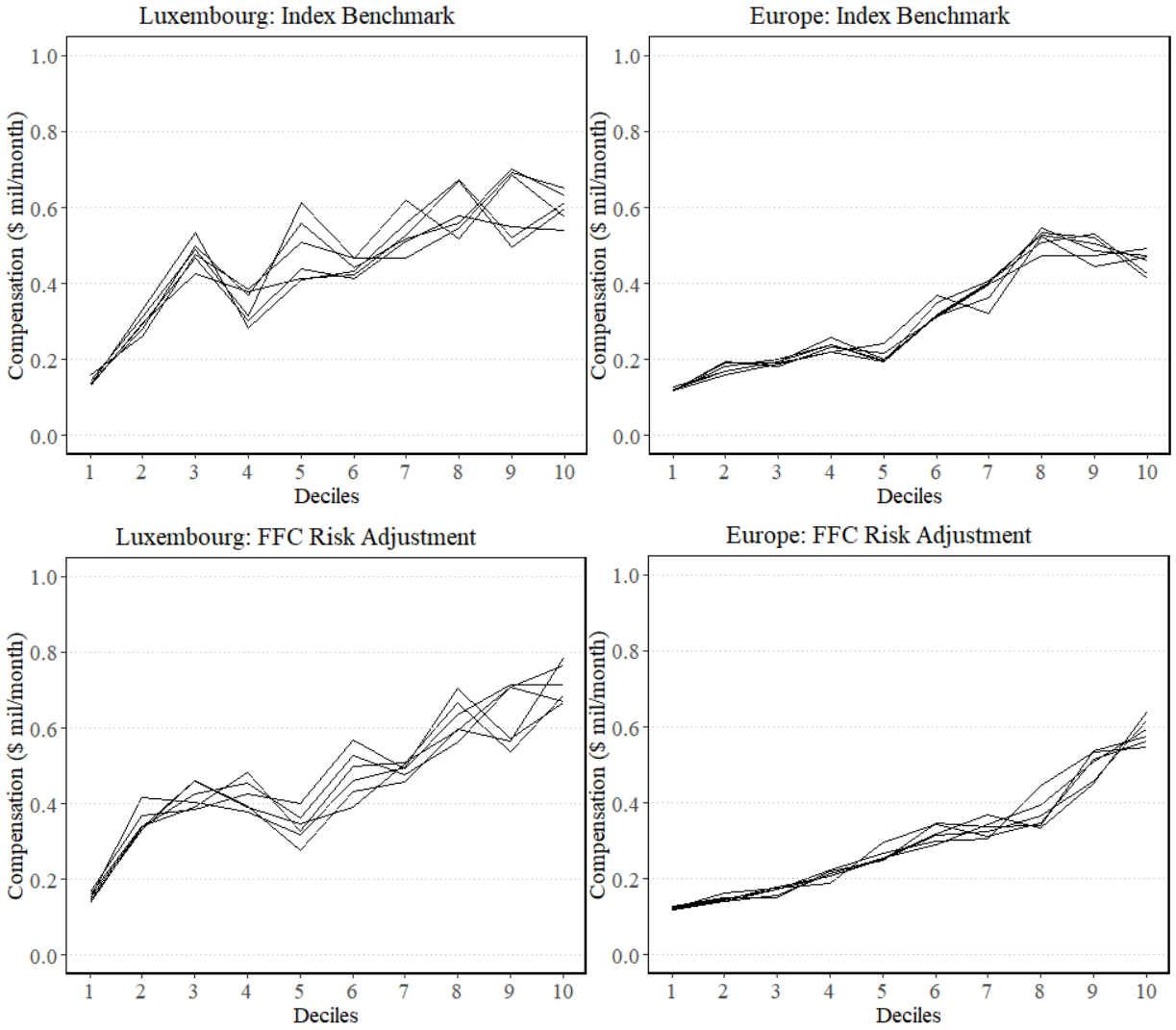

Fig. 5. Out-of-sample monthly compensation for Luxembourg and Europe: The figure displays the skill-ratio sorted average out-of-sample monthly compensation for each decile for the Vanguard Benchmark and FFC factor specification. Each line represents the horizons of 3 to 8 years. The figure shows the result for Luxembourg and Europe.

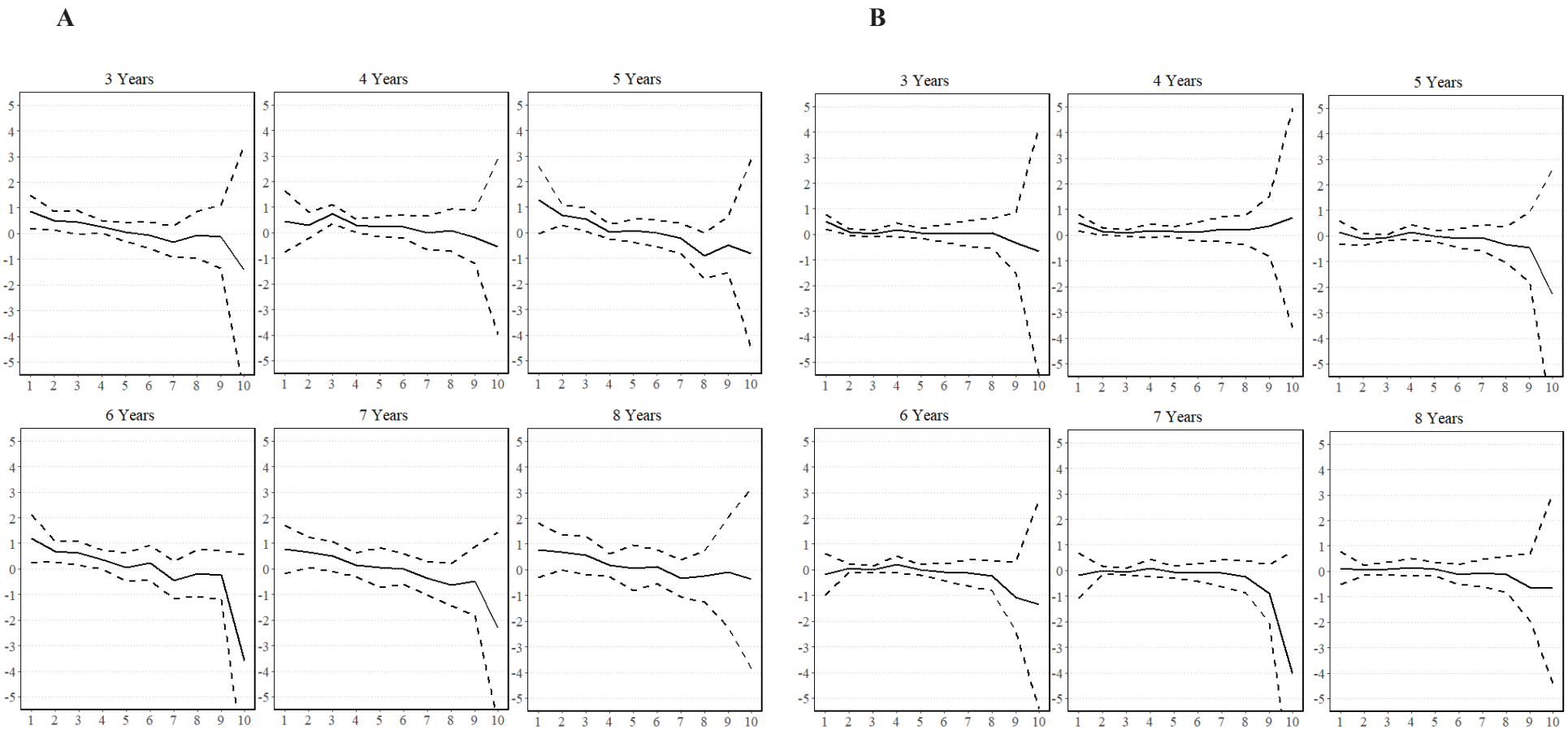

Fig. 6. Sorting index benchmark derived value-added on compensation for Luxembourg and Europe: The figure displays out-of-sample value-added, $\hat{s}_i$, for each fund in millions of Y2000 dollars/month extracted using the index benchmark, sorted into deciles based on compensation (the solid line). Each graph shows the results for the different measurement horizons indicated by the title, and the deciles are indicated by the horizontal axes. The dotted lines indicate the two standard error bounds. Panel A shows the results for Luxembourg funds, and panel B shows the results for European funds.

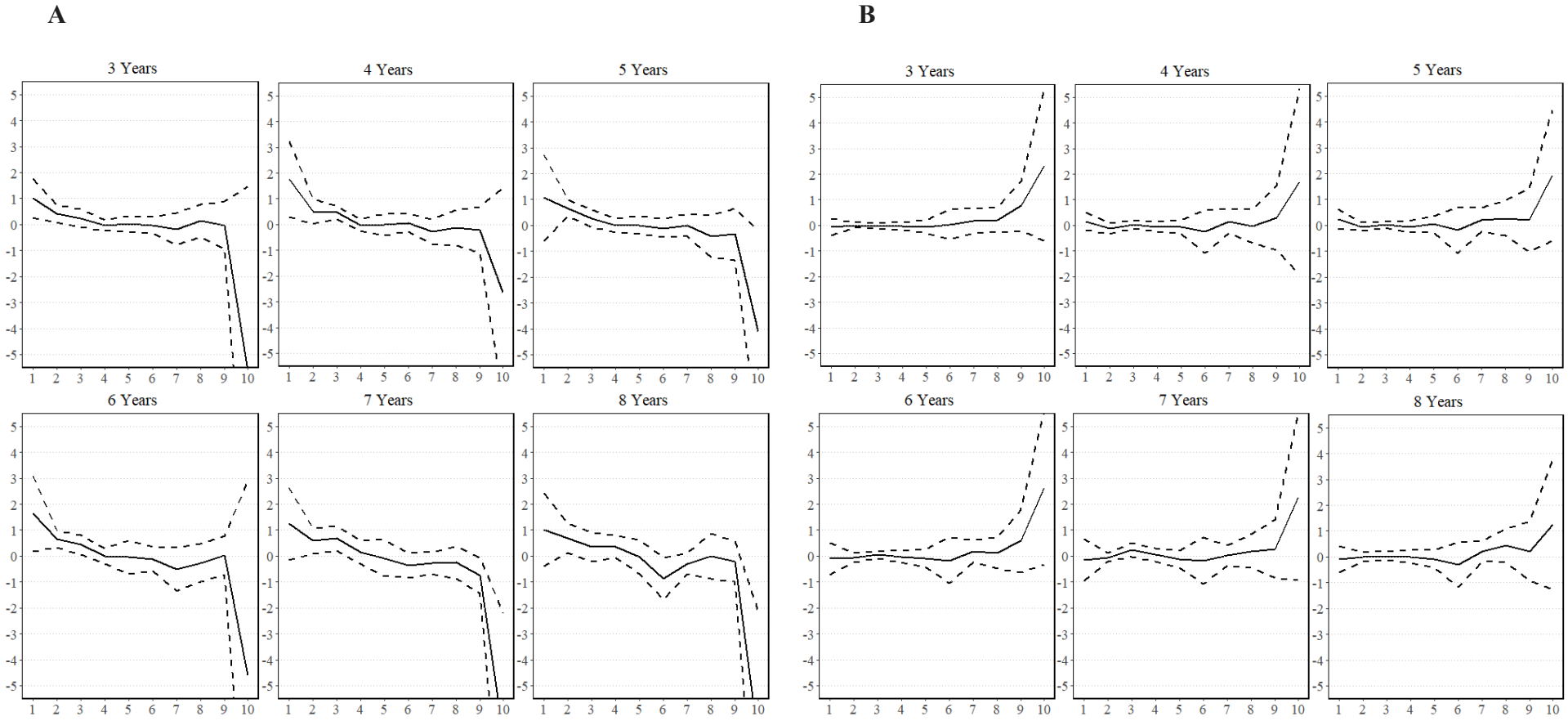

Fig. 7. Sorting FFC factor specification derived value-added on compensation for Luxembourg and Europe: The figure displays out-of-sample value-added, $\widehat{s}_i$, for each fund in millions of Y2000 dollars/month extracted using the FFC factor specification, sorted into deciles based on compensation (the solid line). Each graph shows the results for the different measurement horizons indicated by the title, and the deciles are indicated by the horizontal axes. The dotted lines indicate the two standard error bounds. Panel A shows the results for Luxembourg funds, and panel B shows the results for European funds.

7. Conclusion

In this study, we use the value-added approach to assess the managerial skill of fund managers in Luxembourg's mutual fund market. Despite our initial finding of both Luxembourg and Europe creating value, overstated t-statistics made us suspect these results. Thus, a persistence test was performed to allay potential concerns. However, in contrast to Berk and van Binsbergen (2015), we do not find persistence of managerial performance in neither Luxembourg nor Europe. The top-decile managers do not generate any value of significance and at times even destroy it. Thereby, our results challenge the notion of persistence in managerial performance among Luxembourg and European fund managers. Investors, presented by Berk and van Binsbergen (2015) as better separators of skill, actually seem to predict skill worse, even undercompensating managers who subsequently created value. These findings contradict our hypothesis and suggest that managers in the mutual fund market of Luxembourg do not outperform the benchmark.

8. References

Baik, Bok, Jun-Koo Kang, and Jin-Mo Kim. 2010. Local Institutional Investors, Information Asymmetries, and Equity Returns. Journal of Financial Economics, 97 (1): 81–106. https://doi.org/https://doi.org/10.1016/j.jfineco.2010.03.006

Berk, J., & Green, R. 2004. Mutual Fund Flows and Performance in Rational Markets. Journal of Political Economy, 112(6), 1269–1295. https://doi.org/10.1086/424739

Berk, J. B., & van Binsbergen, J. H. 2015. Measuring skill in the mutual fund industry. Journal of Financial Economics, 118(1), 1–20. https://doi.org/https://doi.org/10.1016/j.jfineco.2015.05.002

Bollen, N. P. B., & Busse, J. A. 2001. On the Timing Ability of Mutual Fund Managers. The Journal of Finance, 56(3), 1075–1094. http://www.jstor.org/stable/222543

Carhart, M. M. 1997. On Persistence in Mutual Fund Performance. The Journal of Finance, 52(1), 57–82. https://doi.org/10.2307/2329556

Christoffersen, Susan E K, and Sergei Sarkissian. 2009. City Size and Fund Performance. Journal of Financial Economics, 92 (2): 252–75. https://doi.org/https://doi.org/10.1016/j.jfineco.2008.06.001.

Coval, Joshua D, and Tobias J Moskowitz. 2001. The Geography of Investment: Informed Trading and Asset Prices. Journal of Political Economy, 109 (4): 811–41. https://doi.org/10.1086/322088.

Fama, E. F., and French, K. R. 1995. Size and Book-to-Market Factors in Earnings and Returns. The Journal of Finance, *50*(1), 131–155. https://doi.org/10.2307/2329241

Fama, E. F., and French, K. R. 2010. Luck versus Skill in the Cross-Section of Mutual Fund Returns. The Journal of Finance 65 (5): 1915–47. https://doi.org/https://doi.org/10.1111/j.1540-6261.2010.01598.x.

Gruber, M. J. 1996. Another Puzzle: The Growth in Actively Managed Mutual Funds. The Journal of Finance, 51(3), 783–810. https://doi.org/10.2307/2329222

Kaniel, R., Lin, Z., Pelger, M., & van Nieuwerburgh, S. 2023. Machine-learning the skill of mutual fund managers. Journal of Financial Economics, 150(1), 94–138. https://doi.org/https://doi.org/10.1016/j.jfineco.2023.07.004

Kosowski, R., Timmermann, A., Wermers, R., & White, H. 2006. Can Mutual Fund “Stars” Really Pick Stocks? New Evidence from a Bootstrap Analysis. The Journal of Finance, 61(6), 2551–2595. http://www.jstor.org.ez.hhs.se/stable/4123439

9. Appendix A: Tables and Figures

Table 1A: Index Names

The table shows a list of index words that are used in order to identify the index funds from the raw data set.

DAX	Access	Nasdaq 100
Stoxx	Eurostoxx	OMXC20
FI	Indeksi	Euro 50
Euro countries 50	Indiciel	Insights Equity
Indx	Zurich Carbon Neutral World Equity	OMX
MSCI	S&P 500	OMXS30
Índice	UK 100 companies	ECPI
Indexed	Idx	Passive
Track	IEF	Passiv
Index	CAC 40	Dow Jones
FTSE	Europe 50	Indice
Indeks		

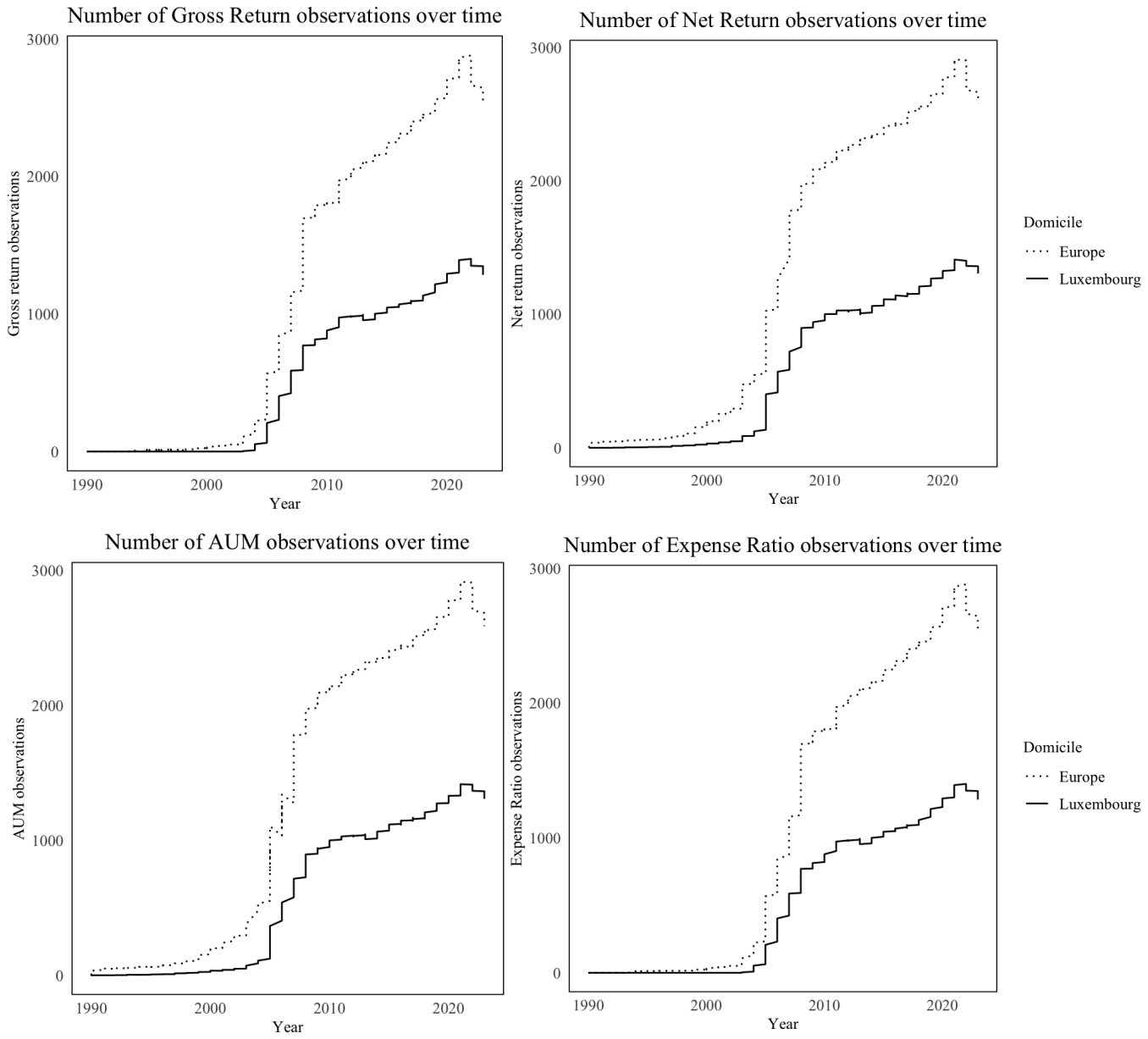

Fig. 1A. Number of observations over time for Luxembourg and Europe: The figure shows the number of observations per month of monthly net return, monthly gross return, monthly AUM, and monthly expense ratio variables for the Luxembourg and European data sets.

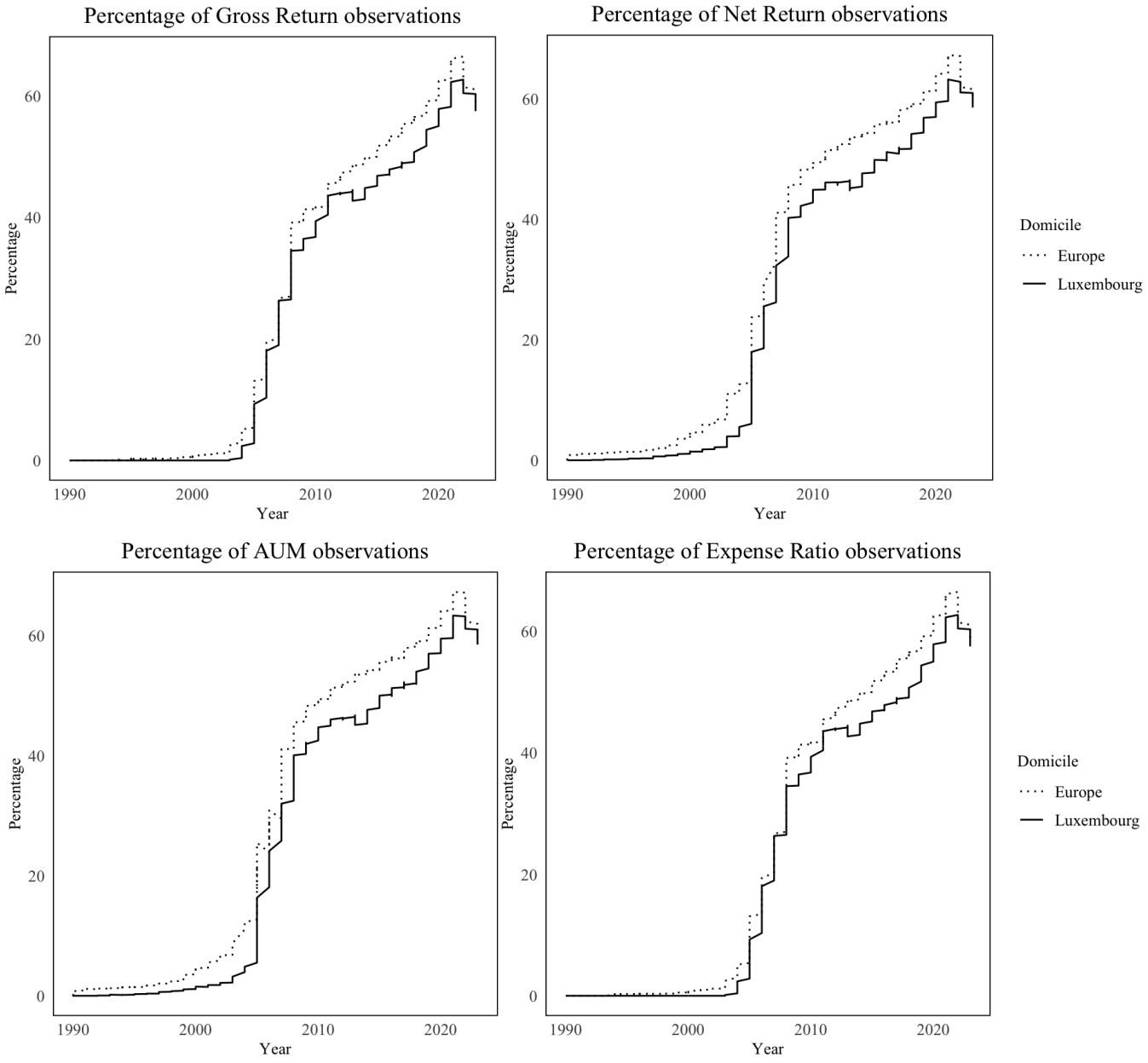

Fig. 2A. Percentage of total observations over time for Luxembourg and Europe: The figure shows the percentage of the number of funds reporting the monthly gross return, monthly net return, monthly AUM, and monthly expense ratio at a certain point in time for the total number of funds of the total data set.

.

10. Appendix B: Additional data information

10.1 Extracted variables

To extract equity mutual funds we set the search criteria to drop all observations that have less than 75% equity by using the *Asset alloc equity % (Net)* variable. Additionally, we included non surviving funds over the time period. Based on these conditions we downloaded the time series data of *monthly return in base currency* (net return), *Net Asset per share class in USD* (AUM), and *monthly gross return in base currency* (gross return). It is important that the AUM is downloaded in USD because it avoids intra-European currency comparison problems and enables inflation adjustment through the Consumer Price Index. Additionally, the return data is downloaded in base currency and not in USD as foreign exchange-rate risk does not become a problem then. If not accounted for the excess return will be affected, for which the chosen interest rate is described in section 10.3.

10.2 Inflation adjustment

The inflation adjustment was enabled through the Consumer Price Index (CPI) retrieved from *Wharton Research Data Service*. CPI was given as a change in consumer price index and therefore to be able to inflation adjust the AUM, the cumulative product is entailed. This required a conversion of the change of rate to $1+r$ format by utilising January 2000 as the basis year. From this, we divided the AUM at time t with the cumulative product of the index.

10.3 Interest rate

To compute excess returns for our data we need a proxy for the global interest rate. We determined this proxy as the 10-year yield to maturity U.S. T-bill, based on the prominence and liquidity of the U.S. government securities market. Given that treasury bills are highly liquid and actively traded, they reflect market expectations and reflect changes in economic conditions. The subsequent T-bill was downloaded from *S&P Capital IQ*.