DEUTSCHE
BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 28/2021

Return differences between DAX ETFs and the benchmark DAX

Christoph Schmidhammer
(Deutsche Bundesbank University of Applied Sciences)

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

ISBN 978–3–95729–836–2
ISSN 2749–2958

Non-technical summary

Research question

Exchange Traded Funds (ETFs) have become important products since they can be traded continuously during the day and costs of passive investment strategies are low. Annual financial statements are published by the German Federal Gazette and include relevant costs and income information of funds. However, this is not satisfactory for long-term investors since information is provided only once a year at a discrete point in time. The purpose of this paper is to continuously identify annual cost and income of passively managed products.

Contribution

Although cost and income information is not available daily, an approximation is possible due to changes in net asset values (NAVs) which represent fair values of a fund, determined each trading day. While costs lead to a decrease in NAVs relative to its benchmark, income components lead to an increase. To identify costs and income, this study compares NAV returns with benchmark returns. An analysis of long-term costs is implemented by the construction of a daily rolling window of annual return differences between ETFs and the corresponding index. The generated time series allows me to compare product specific costs and to identify relevant drivers such as returns of Germany's leading equity index (DAX) and market makers.

Results

For the DAX index market, annual return differences considerably exceed total expense ratios as publically available cost information. Product-specific return differences are significant, however, differences between DAX ETFs and the DAX index tend to converge over time. For all DAX ETFs, return differences are significantly influenced by DAX index returns and by market maker prices. Product characteristics such as cash holdings and use of profits deliver valuable arguments to explain these findings.

Nichttechnische Zusammenfassung

Fragestellung

Wichtige Erfolgsfaktoren für börsengehandelte Fonds, sogenannte Exchange Traded Funds (ETFs), sind neben dem Intraday-Handel auch die geringen Gebühren. Eine genaue Zusammensetzung von Kosten und Erträgen, die ein ETF während eines Geschäftsjahres verursacht und erwirtschaftet, wird im Bundesanzeiger veröffentlicht. Dennoch ist die Informationslage aus der Perspektive von Langzeitinvestoren nicht zufriedenstellend, da Kosten nur einmal pro Jahr stichtagsbezogen vorliegen. Ziel der Arbeit ist es, jährliche Kosten- und Ertragsstrukturen passiver Indexprodukte fortlaufend zu identifizieren.

Beitrag

Obwohl Kosten- und Ertragspositionen nicht täglich bekannt sind, gelingt deren Identifikation implizit durch die Auswirkung auf die Höhe des täglich verfügbaren fairen Fondswerts, den Nettoinventarwert (NAV). Während Kosten den NAV reduzieren, erhöhen Erträge den fairen Fondswert. Beides wirkt sich unmittelbar auf die NAV-Rendite aus. Ein Renditevergleich zwischen ETF-Renditen und Indexrenditen lässt somit Rückschlüsse auf die aufgelaufenen Kosten und Erträge zu. Die jährlichen Renditedifferenzen werden täglich rollierend ermittelt und bilden die Basis für die Analyse. Die so generierte Datenreihe ermöglicht einen aussagekräftigen Vergleich produktspezifischer Kosten und Erträge sowie die Identifikation relevanter Einflussfaktoren wie beispielsweise Renditen des Deutschen Aktienindex (DAX) und Market-Maker-Preise.

Ergebnisse

Für den DAX-Indexmarkt lässt sich feststellen, dass die jährlichen Renditedifferenzen von DAX-ETFs zum DAX deutlich höher ausfallen als die veröffentlichten Kosten (Total Expense Ratios). Die Kosten zwischen den ETFs unterscheiden sich signifikant, die Unterschiede nehmen jedoch im Zeitablauf deutlich ab. DAX-Renditen und Market-Maker-Preise üben einen signifikanten Einfluss auf Renditedifferenzen zwischen DAX-ETFs und dem DAX aus. Die Ergebnisse lassen sich durch Produkteigenschaften wie Barreserve oder Dividendenverwendung erklären.

Deutsche Bundesbank Discussion Paper No 28/2021

Return differences between DAX ETFs and the benchmark DAX

Christoph Schmidhammer*

Deutsche Bundesbank University of Applied Sciences

Abstract

For the DAX index market, this paper analyses the development of return differences between exchange traded funds (ETFs) and the DAX index from the perspective of long-term investors. The newly introduced methodology provides the opportunity to continuously identify long-term costs of passively managed products independent from the information of annual financial statements. This enables to test for product-specific return differences and to identify relevant cost drivers such as index returns and market makers. Results reveal that on average, DAX ETFs costs considerably exceed total expense ratios. Product-specific return differences are significant, however, differences tend to converge over time. For all ETFs, deviations are significantly influenced by index returns. Product characteristics deliver valuable arguments to explain these findings. Also market makers significantly contribute to return differences.

JEL-Classification: G12, G13, G14

Keywords: Exchange Traded Funds, Net Asset Value, market maker prices, return differences, Total Expense Ratio, ETF issuers, rolling window

* Contact address: Deutsche Bundesbank University of Applied Sciences, Schloss, 57527 Hachenburg, Germany. Phone: +49 2662 83 545. E-Mail: Christoph.schmidhammer@bundesbank.de. This paper has benefited from helpful comments and suggestions by Christian Walkshäusl, Christoph Memmel, Andreas Kermer, Christopher Priberny and Martin Schmidhammer. The views expressed in this paper are those of the author and do not necessarily coincide with the views of the Deutsche Bundesbank or the Eurosystem.

1. Introduction and literature review

Exchange Traded Funds (ETFs) have become important products for professional investors and private households. As passively managed products, ETFs seek to replicate an underlying benchmark. ETFs can be continuously traded during the day, prices are publicly available and costs of passive investment strategies are low. According to Ben-David, Franzoni and Moussawi (2017), such characteristics substantially contribute to the popularity of ETFs.[1] Professional investors frequently use ETFs for hedging and arbitrage strategies, where the focus is on short-term periods such as intraday time intervals. In this context, a body of studies analyses the replication quality of ETFs and price innovation dynamics.[2]

However, the perspective of long-term investments provides an open field for research. Especially during a period of low interest rates, long-term investments in well-diversified portfolios such as index ETFs represent an interesting opportunity. Passively managed products tremendously increased during the last years and a high diversity of products emerged. For comparable products in highly liquid and transparent markets, characteristics such as the cost structure can be relevant for long-term investment decisions. Charupat and Miu (2013), for example, state that long-term performance of ETFs is significantly affected by deviations of ETF and index returns. However, long-term cost information is rare. Although annual financial statements are published, e.g., by the German Federal Gazette[3] and annual total expense ratios (TER) which summarize all cost components as described in Dorfleitner, Gerl and Gerer (2018) are announced, the information is provided only once a year at a discrete point in time. Since

[1] Following Markowitz (1952) and Tobin (1958), financial theory delivers further arguments for the success of ETFs, since the risk-return relation of well-diversified portfolios is expected to be efficient.

[2] The regression specification as illustrated in Sharpe (1964) can be applied to analyse the replication quality of index products. Besides the replication quality of ETFs, Dorfleitner, Gerl and Gerer (2018) provide an interesting analysis of the pricing efficiency of exchange-traded commodities. ETF arbitrage is, for example, studied in Richie, Daigler and Gleason (2008), Cherry (2004) and Petajisto (2017). Hedging opportunities are addressed in Alexander and Barbosa (2008). Hasbrouck (2003) and Chen, Lin, Cou and Lang (2002) discuss ETF-related price discovery.

[3] The German Federal Gazette is issued by the Federal Ministry of Justice and provides open accessible information on corporate and financial announcements.

long-term investment decisions are not bound to a specific date, a continuous analysis of long-term costs can provide valuable information for investors. This paper aims to fill that gap by analysing annual costs on a day-to-day level.

First, a suitable proxy is necessary to continuously identify long-term costs, e.g., on daily level. Passively managed products that seek to replicate an underlying benchmark provide the advantage, to be easily compared to its reference. For ETFs, net asset values (NAVs) which represent fair values of a fund can be continuously compared with its benchmark, since NAVs are determined each trading day.[4] Costs such as transaction costs and management fees are promptly reflected in NAVs. While costs lead to a decrease in NAVs relative to its benchmark, income components such securities lending lead to an increase. To approximate costs, this study compares NAV returns with benchmark returns. This is necessary since NAVs do not one-on-one correspond to the benchmark, e.g., due to management fees. Second, a day-to-day analysis of long-term costs can be implemented by the construction of a daily rolling window of long-term return differences between passively managed products and the corresponding benchmark. I rely on Ghysels, Santa-Clara and Valkanov (2005) who state that the construction of rolling windows requires an optimal window size. To address the perspective of long-term investors and to fully capture cost and income components that occur during one business year, annual periods are used in this study. This is also in line with annual financial statements as described in Baltussen, Post and Vliet (2012). In addition, results can be compared to annual cost information such as TER.

The purpose of this paper is to analyse costs of ETFs, that replicate the DAX, Germany's main equity index from the perspective of long-term investors. Relying on Elton, Gruber and Busse (2004), prices of products with identical cashflows are expected to converge for highly

[4] NAVs include actual values of a fund's total assets and cash components.

liquid and transparent markets. Since this is the case for the DAX index market, I expect costs of DAX ETFs to converge.[5] In addition to the hypothesis of efficient markets, factors that might influence deviations between DAX ETFs and the DAX index such as market makers, index returns and turnover are analysed. A daily rolling window of annual return differences between NAV returns and DAX returns reveals that on average, costs range between 0.234% and 0.699% and considerably exceed TERs. According to Schmidhammer (2014), NAV results are of special interest for institutional investors who trade over the counter at fair values. Since private investors cannot trade over the counter, closing prices (CPs) which can be attributed to market makers are included. On average, CP related costs range between 0.230% and 0.701% and are comparable to NAV results. Regression results reveal that DAX index returns significantly contribute to return differences. Product characteristics as illustrated in table 1 deliver valuable arguments to explain the findings. It is interesting to observe that product-specific return differences strongly decrease over time. Also market makers significantly drive return differences.

In the financial literature, the replication quality of passive investment products is an important field of research. Roll (1992), for example, describes the tracking error as a measure of passive management performance. A low tracking error implies that the replication quality is high, which leads to a low risk of a passive investment product significantly underperforming a benchmark. Rompotis (2006) examines the tracking error of US and international ETFs. The author shows that ETFs move in line with the underlying benchmark. Cremers and Petajisto (2009) analyse the success of active management by controlling for the influence of tracking errors. Johnson (2009) explains differences in tracking error values between foreign index

[5] The selected products can be continuously traded during the day. Börse Frankfurt, for example, provides information about the historic development of an ETF as well as most recent trading prices. Xetra liquidity measures of analysed ETFs (also provided by Börse Frankfurt) closely correspond to the liquidity measures of DAX 30 shares.

ETFs. For the early 2000s, Lang and Röder (2008) provide a detailed analysis of the DAX EX, Germany's first ETF replicating the DAX index. Their results show a high replication quality of the DAX EX. This is in line with Kundisch and Klein (2009), who observe a low tracking error of the DAX EX compared to DAX index funds. Rompotis (2012) analyses 43 ETFs between 2003 and 2005. He finds that the tracking error positively corresponds to risk, premium and spread. Merz (2015) identifies shortcomings of tracking error measures, since deviations from the benchmark are not sufficiently captured, for example, when management fees are proportionally deducted from assets under management (AUM). To overcome shortcomings of tracking error measures, the present study analyses return differences between DAX ETFs and the DAX index.[6] This allows costs and income to be accounted for since negative and positive outcomes are possible. Negative differences between DAX ETF returns and DAX index returns can be interpreted as costs and positive differences as income.[7]

Further studies focus on cost components. Anderson and Highfield (2014) address annual management fees captured by expense ratios. Including more than 1,100 ETFs in 2012, the authors find that fees vary between 0.32% and 0.92% for portfolio subsamples. On average, the annual expense ratio amounts to 0.61%. The authors also analyse factors that affect expense ratios, such as turnover. In the context of mutual funds, LaPlante (2001) illustrates that stock fund expense ratios slightly decrease for a sample during 1994 to 1998. Their study controls for factors such as market volume, age, year and index. Dorfleitner, Gerl and Gerer (2018) analyse the pricing efficiency of exchange-traded commodities. Besides transparent management fees, the authors describe a variety of components that potentially influence total costs of exchange-traded commodities. Elton et al. (2002) illustrate the diversity of cost components for SPDRs.

[6] Return differences between index products and the underlying index are also applied in Osterhoff and Kaserer (2016). However, absolute values are determined as tracking error measure.
[7] Johnson (2009), for example, applies annual return differences between foreign country index returns and US index returns to explain tracking error values.

Examples are management fees, transaction costs, dividends not reinvested in the fund and factors that affect the replication quality. For ETFs, the total expense ratio (TER) is expected to capture all of the costs.[8] However, this is not necessarily the case, as diverse studies show. Elton, Gruber and Busse (2004) identify return differences of up to 2.09 percent between S&P 500 index funds, which are not solely explained by annual expense ratios. More recently, Blitz, Huij and Swinkels (2012) uncover annual return differences between 50 and 150 basis points for European index funds and ETFs. Substantial parts of the underperformance are not captured by the TER. Lang and Röder (2008) analyse cost components of the DAX EX between 2001 and 2005. The authors rely on a detailed analysis of issuers' annual financial statements.

However, a continuous development of cost and income structures beyond annual financial statements as applied in this study has not yet been addressed. The construction of a daily rolling window of annual DAX ETF returns compared to annual DAX index returns[9] allows (1) to identify long-term costs and income beyond annual financial statements, (2) to identify long-term costs and income on a daily level and (3) to apply statistical tests to capture product specifications. In line with the hypothesis of Elton, Gruber and Busse (2004), for DAX ETFs, product-specific return differences converge over time. In contrast to the prominent study by Gil-Bazo and Ruiz-Verdú (2009), who observe a negative dependence between the performance and the fees of mutual funds, negative annual return differences between DAX ETFs and the DAX index (which can be interpreted as costs) increase when DAX returns increase. An explanation can be found in issuer's management of cash holdings and use of profits.[10]

[8] For DAX ETFs, TER definition and TER values can be found in the German Federal Gazette. Costs relate to separate assets.
[9] Rolling windows are applied, for example, in prominent studies such as Merton (1980) or French, Schwert and Stambaugh (1987).
[10] Also regulatory requirements relating to undertakings for collective investment in transferable securities (UCITS) determine cash holdings.

Besides NAVs, the inclusion of CPs serves as a robustness test and makes it possible to identify the influence market makers have on return differences. Engle and Sarkar (2006) and Petajisto (2017) identify deviations between ETF prices and NAVs, which supports the assumption that market makers influence costs. The present study shows that market makers significantly influence costs. The results of the analysis remain robust when different AR terms are included or Newey and West (1987) is employed to correct for heteroscedasticity and autocorrelations in the sample's residuals. Further robustness tests are illustrated in section 4.

The remainder of the paper is structured as follows. Section 2 describes the data and illustrates the construction of a daily rolling window of annual return differences. Section 3 presents the methodology and empirical results. Section 4 illustrates robustness tests. Section 5 summarises the results and concludes.

2. Data

The sample is based on daily observations of DAX values, DAX ETF prices and NAVs. Since NAVs are determined at the end of each trading day, the corresponding closing values of the DAX and CPs of DAX ETFs are applied. The DAX index market includes eleven DAX ETFs. For the analysis, only products are selected that provide an adequate time series and comparable product characteristics. Therefore, all seven DAX ETFs issued before 2016 representing a fraction of 1/100 of the DAX index are included. The sample covers a broad market spectrum that includes all market leading DAX ETFs as well as products with a small market share. ETFs can be continuously traded during the day and the duration is unlimited. CPs, represented by most recent traded prices and DAX values are from Bloomberg.[11] NAVs, represented by last indicative net asset values are from Bloomberg, Thomson Reuters Eikon and DAX ETF issuers.

[11] Source: Bloomberg Finance L.P.

Table 1: Product characteristics of DAX ETFs

Product	Issuance	Deposit	No. Market Makers	Use of Profits	Replication Method	TER
ETF 1	2001	GER	11	acc.	full	0.16%
ETF 2	2007	LU	10	acc.	full (swap)	0.09%
ETF 3	2008	GER	4	acc.	full	0.15%
ETF 4	2006	LU	8	acc.	full	0.15%
ETF 5	2008	LU	4	distr.	full (swap)	0.08%
ETF 6	2012	LU	8	distr.	full	0.09%
ETF 7	2015	GER	4	distr.	full	0.15%

Table 1 illustrates product characteristics of the selected DAX ETFs such as deposit domicile (Germany = GER, Luxembourg = LU), number of competing market makers, use of profits (accumulating versus distributing), replication method (full replication versus swap based) and TER.[12] The deposit domicile is important to account for different tax regimes. In order to make ETFs comparable, the use of profits is further considered. While ETFs 1 to 4 are accumulating, dividends of ETFs 5 to 7 are distributed. Table 1 also indicates changes from a swap-based to a full replication method for ETFs 2 and 5. For all DAX ETFs, annual fees range between 0.08% and 0.16%. However, TER does not capture all annual costs, e.g. transaction costs, institutional custodian fees, account maintenance charges, IT information services and costs of enforcing legal claims are not included.[13] While costs lead to a decrease in a fund's value, income components such as dividends, securities lending and interest lead to an increase.[14] Cost components as well as income components are disclosed annually. From the perspective of continuous long-term investor decisions, financial statements comprising cost and income information once a year are not satisfactory. The purpose of this paper is to uncover

[12] Product characteristics are from Deutsche Boerse AG (January 2020). Possible changes of market makers over time are not captured.
[13] Annual cost information stems from the German Federal Gazette.
[14] Income components are illustrated in the German Federal Gazette.

daily patterns of long-term investments. Addressing the purpose of passively managed products, I compare annual DAX ETF returns with annual DAX index returns as a benchmark.

The inclusion of NAVs allows to address the perspective of institutional investors. Since ETFs are not traded at fair value, the present study also includes CPs to address the perspective of private investors. This provides another opportunity to identify the influence market makers have on costs and serves as a robustness test for NAV results. Equation (1) shows the daily calculation of annual return differences, denominated as a rolling window:

$$R_{i,t}^{diffNAV} = R_{i,t}^{NAV} - R_t^{DAX} \tag{1}$$

Annual return differences of NAV i ($i = 1, \ldots\ldots, 7$) on day t, $R_{i,t}^{diffNAV}$, are calculated by subtracting annual DAX index returns on day t, R_t^{DAX}, from annual NAV returns on day t, $R_{i,t}^{NAV}$. The corresponding return differences of CP i ($i = 1, \ldots\ldots, 7$) on day t, $R_{i,t}^{diffCP}$, are calculated as follows:

$$R_{i,t}^{diffCP} = R_{i,t}^{CP} - R_t^{DAX} \tag{2}$$

Since DAX ETFs are passively managed, DAX index returns and DAX ETF returns are highly correlated. The construction of return differences as dependent variable, where DAX returns are subtracted from NAV and CP returns, eliminates the influence of DAX index returns for $R_{i,t}^{diffNAV}$ and $R_{i,t}^{diffCP}$. The so constructed dependent variables allow for positive and negative outcomes, independent from the development of DAX index returns. In the case, where a product perfectly replicates the benchmark, return differences amount to zero. Negative values can be interpreted as costs and positive values as income. Hence, endogeneity problems do not occur, when DAX returns are included as independent variable in regression specifications.

According to Blitz, Huij and Swinkels (2012), tax regimes significantly influence the performance of passive investment products. To make products comparable, the present study considers differences in tax regimes of the seven DAX ETFs. Relying on table 1 characteristics, accounts of ETFs 1, 3 and 7 are administrated in Germany, while accounts of ETFs 2, 4, 5 and 6 are administrated in Luxembourg. For German accounts, issuers are required to pay taxes and solidarity surcharges. Hence, return differences of ETFs 1, 3 and 7 are adjusted by an annual amount of taxes and solidarity surcharges relative to assets under management. The information on tax payments and solidarity surcharges is publicly available in the German Federal Gazette. Tax payments due in Luxembourg are also publicly available and return differences are adjusted correspondingly.[15] For ETFs 1 and 3, tax payments are disclosed on an overall fund level. To adjust $R_{i,t}^{diffNAV}$ and $R_{i,t}^{diffCP}$, tax payments are divided by the fund's total assets. Relations are added to $R_{i,t}^{diffNAV}$ and $R_{i,t}^{diffCP}$. For ETFs 2, 4, 5, 6 and 7, payments are disclosed per share and can be directly added to NAVs and CPs, before $R_{i,t}^{diffNAV}$ and $R_{i,t}^{diffCP}$ are determined.

To establish an unbiased analysis of costs between products, the use of profits is also considered. In the case of accumulating funds, dividends are reinvested in the fund. In the case of distributing funds, dividends are paid out which leads to decreases in NAVs and CPs of distributing funds relative to accumulating funds. Since dividend payments and dates are publicly available, NAVs and CPs of distributing funds are adjusted by the amount of dividend payments.[16] The chosen methodology provides the opportunity to make distributing and accumulating funds comparable. For investors, dividend payments represent an increase in cash

[15] Tax payments due in Luxembourg are published, e.g., by the German Federal Gazette.
[16] Dividend payments and dates are published by issuers.

(which corresponds to an adjustment of NAVs and CPs). This contributes to explain the product specific influence of DAX index returns as shown, for example, in table 6.

Table 2 illustrates descriptive statistics of annual DAX returns, adjusted return differences between NAVs and the DAX, and adjusted return differences between CPs and the DAX.[17] The time span of the study is January 2010 to December 2018. Due to the calculation of annual returns, the total sample includes data from January 2009 to December 2018. Outliers that exceed +/-4% are excluded from the sample. Since ETFs 6 and 7 are issued in 2012 and 2015 respectively, the number of observations lies below ETFs 1 to 5. While a ten year sample of discrete annual returns provides ten observations per product only, the application of a daily rolling window of annual returns significantly increases the number of observations (2,281 as a maximum). This allows me to test for product specifications as applied in OLS regressions. However, autoregressive disturbances have to be considered.

For the DAX index, the sample period includes upwards and downwards movements between -21.752% as a minimum and 60.263% as a maximum. On average, the DAX increased by an annual mean of 11.570%. Return differences based on NAVs range from -0.234% for ETF 5 to -0.699% for ETF 6. Negative values indicate that annual product returns are below annual DAX index returns, which represent costs for long-term investors. These results are interesting since the TERs of the seven ETFs range between 0.08% and 0.16% p.a. Although return differences are qualitatively low, they considerably exceed TERs. Further analysis shows that high negative values of ETFs 6 and 7 are strongly determined by dividend payments.

Since ETFs are not traded at fair value, market-maker CPs are included in the analysis. In addition to robustness purposes, this makes it possible to identify the influence market makers have on costs. With a range of -0.230% for ETF 3 to -0.701% for ETF 6, average CP

[17] Adjustments concern taxes and solidarity surcharges for German deposit accounts.

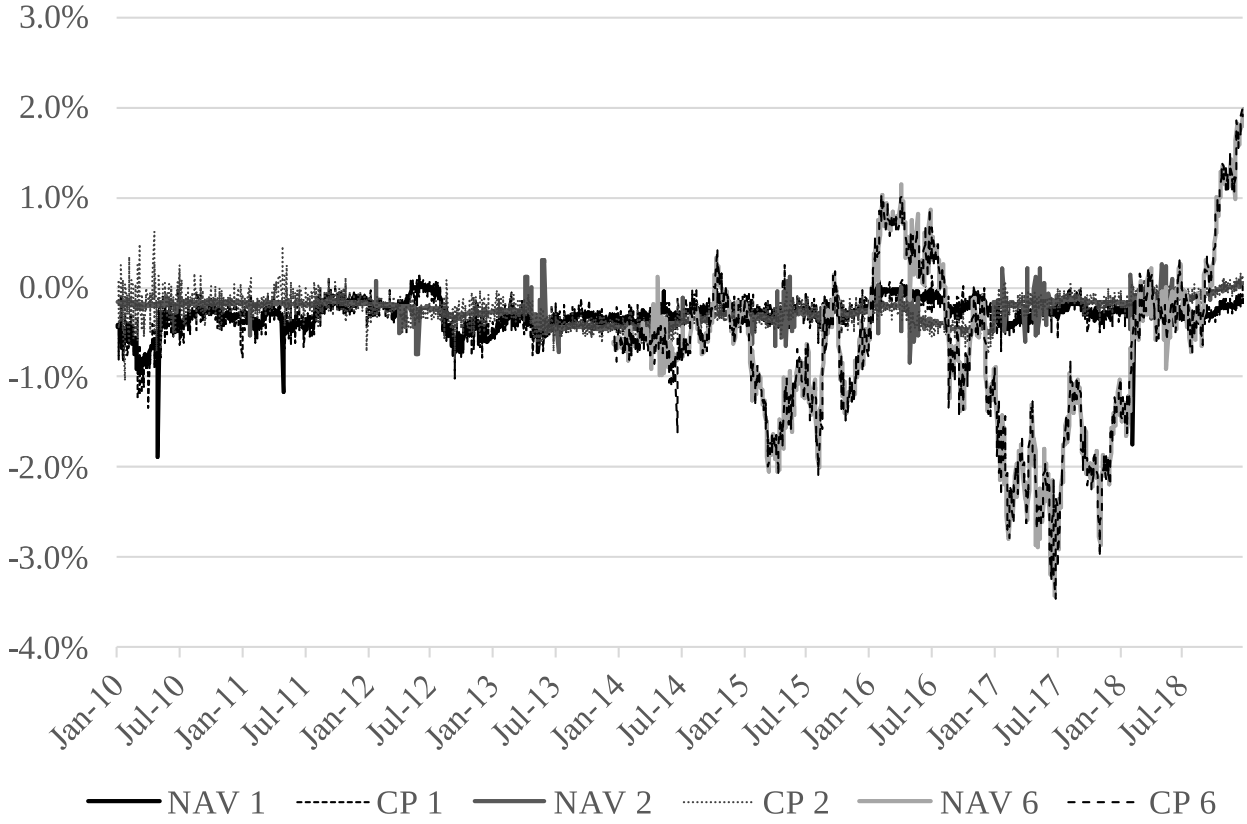
deviations largely correspond to the return differences of NAVs. However, differences between NAVs and CPs can be observed which can be interpreted as the market maker's responsibility.

Table 2: Annual returns (%) of the DAX, annual return differences (%) between the DAX and NAVs, and between the DAX and CPs from January 2010 to December 2018

	DAX	ETF 1	ETF 2	ETF 3	ETF 4	ETF 5	ETF 6	ETF 7
				NAVs				
Min.	-21.752	-1.884	-0.831	-2.150	-2.753	-3.585	-3.422	-1.620
Max.	60.263	0.041	0.311	0.131	1.490	3.257	1.992	2.159
Mean	11.570	-0.290	-0.252	-0.241	-0.363	-0.234	-0.699	-0.450
Median	15.143	-0.290	-0.216	-0.241	-0.351	-0.188	-0.607	-0.541
Std.dev.	15.091	0.148	0.135	0.172	0.149	0.356	0.942	0.445
Observations	2,281	2,281	2,281	2,281	2,281	2,277	1,267	524
				CPs				
Min.		-1.734	-1.033	-3.068	-2.716	-2.124	-3.462	-1.439
Max.		0.250	0.620	2.457	0.538	1.736	1.996	0.926
Mean		-0.288	-0.248	-0.230	-0.359	-0.237	-0.701	-0.468
Median		-0.274	-0.244	-0.237	-0.350	-0.199	-0.630	-0.547
Std.dev.		0.171	0.162	0.273	0.186	0.258	0.941	0.398
Observations		2,281	2,281	2,281	2,281	2,277	1,267	524

Figure 1 illustrates the annual cost structures of ETFs 1, 2 and 6 between January 2010 and December 2018. Solid lines represent annual return differences of NAVs 1, 2 and 6, and dotted lines of market-maker CPs 1, 2 and 6. Products are selected to capture different characteristics such as German versus Luxembourg deposits and accumulating versus distributing funds. Since ETF 6 was issued in November 2012, annual return differences are available from November 2013. Return differences of ETF 6 considerably vary over time for NAVs and CPs. Further analysis shows that variations can be attributed to dividend distribution. For ETFs 1 and 2 (German versus Luxembourg accounts) it can be observed that return differences vary slightly over time. For all products, return differences of market-maker CPs move in line with NAVs.

Figure 1: Annual return differences between the DAX and NAVs 1, 2 and 6, and between the DAX and CPs 1, 2 and 6 from January 2010 to December 2018



3. Methodology and empirical results

This section presents the methodology for analysing return differences between DAX ETFs and the DAX index, and illustrates the results. To address the perspective of long-term investors, annual periods of return differences are selected. This is in line with financial reporting standards as described in Baltussen, Post and Vliet (2012). For annual periods it is assumed to cover total costs necessary to manage a fund. In order to continuously analyse the development of long-term investments, a daily rolling window of annual return differences is constructed. Seminal papers that apply rolling windows are, e.g., Merton (1980) and Ghysels, Santa-Clara and Valkanov (2005). According to Ghysels, Santa-Clara and Valkanov (2005), the length of a rolling window is crucial to identify appropriate risk and return characteristics. Their result delivers a further argument for the present study to determine a rolling window with annual length. More recently, Schmidhammer (2018) analyses a rolling window of bond ladder returns to uncover the performance of bond portfolio strategies. The author uses AR terms since regressions are determined by autoregressive disturbances due to overlapping timeframes. Hence, the present paper also controls for autoregressive (AR) disturbances. Based on the Breusch (1978) and Godfrey (1978) serial correlation LM test, AR terms are applied. For robustness purposes, regression specifications are estimated with different AR terms. Further, Newey and West (1987) is employed to correct for heteroscedasticity and autocorrelations in the sample's residuals. Robustness regressions (not illustrated) qualitatively confirm the results.

To test one central hypothesis as to whether prices of identical products are expected to converge, regression specifications include product dummies. Additionally, factors that affect return differences such as market makers, DAX index returns, DAX return volatility and annual turnover are studied. Daily rolling windows are applied for independent variables and for $R_{i,t}^{diffNAV}$ and $R_{i,t}^{diffCP}$ as dependent variables.

3.1 Regression specifications

The DAX index shows considerable annual upwards and downwards movements between -21.752% and 60.263% (see table 2) during the time span analysed, which leads to comparable movements of related DAX ETFs. In the context of mutual funds, Gil-Bazo and Ruiz-Verdú (2009) discover a negative influence of fund performance on fees.[18] Since DAX ETFs are passively managed, the present paper includes performance components as independent variables such as the risk and return of the related benchmark, the DAX. In contrast to Gil-Bazo

[18] Furthermore, Latzko (1999) and Rompotis (2006) address the impact of fund performance on costs.

and Ruiz-Verdú (2009), due to cash holdings of ETFs a positive relation between costs and DAX index returns is expected. This is the case, since cash does not reflect the development of the underlying index. However, management activities such as securities lending could compensate for increases in negative return differences. Discrete annual DAX returns are measured as the relation between DAX values on day t minus DAX values of the preceding year, $t-1year$, relative to DAX values at time $t-1year$. According to equations (1) and (2), a daily rolling window of annual DAX returns, R_t^{DAX}, is calculated.[19] Following Merton (1980), a rolling window is constructed for return volatility, determined as the standard deviation of annual DAX returns. Consistent with returns, volatility is measured daily and includes annual returns of the preceding year.

Studies such as Malhotra and McLeod (1997), LaPlante (2001) and Rompotis (2012) analyse the influence of turnover on costs. Following Malhotra and McLeod (1997), a turnover ratio is constructed. The present paper calculates the annual turnover for each ETF. The turnover ratio, $Turn_{i,t}$, represents the preceding turnover of ETF i on day t relative to the total turnover of the seven ETFs analysed on day t. Again, consistent with return and return volatility, turnover relies on values of the preceding year.

First, return differences are analysed for NAVs according to equation (1). Since NAVs represent fair values of an ETF, results can be attributed to issuers. To account for annual effects, annual time dummies are included. Equation (3) illustrates the OLS regression specification:

$$R_{i,t}^{diffNAV} = \beta_0 + \beta_1 \cdot R_t^{DAX} + \beta_2 \cdot \sigma_t + \beta_3 \cdot Turn_{i,t} + B' \cdot D_i^{ETF} + C' \cdot D_y^{year} + \varepsilon_{i,t} \qquad (3)$$

[19] The robustness of the econometric approach including dummy variables for DAX index returns as independent variable is addressed in section 4.2.

Return differences $R_{i,t}^{diffNAV}$ of NAV i at time t are defined as dependent variables. The influence of annual DAX index returns R_t^{DAX} on day t is captured by β_1, the effect of the annual DAX return volatility on day t, σ_t , is captured by β_2 and the effect of relative annual turnover, $Turn_{i,t}$, is captured by β_3. Due to different starting points of ETF products, the time series is structured as a panel and product specific effects are captured by dummy variables. B is an $(I-1)\times 1$ vector of coefficients for ETFs i. D_i^{ETF} is the corresponding vector of product dummies. ETF 2 is omitted as the reference.[20] Hence, the constant term β_0 can be interpreted as an ETF-2-specific cost effect. The $(Y-1)\times 1$ vector of yearly time dummies is denoted D_y^{year}. C is an $(Y-1)\times 1$ vector that captures yearly time effects.

Since private investors cannot trade ETFs at fair values, return differences are additionally analysed for CPs. In line with this, $R_{i,t}^{diffCP}$ is defined as a dependent variable. The inclusion of CPs serves as a robustness test and makes it possible to identify the influence of market makers on return deviations. Therefore, differences are determined between $R_{i,t}^{diffCP}$ and $R_{i,t}^{diffNAV}$, denominated as $MM_{i,t}$. Equation (4) illustrates the OLS regression specification:

$$R_{i,t}^{diffCP} = \beta_0 + \beta_1 \cdot R_t^{DAX} + \beta_2 \cdot \sigma_t + \beta_3 \cdot Turn_{i,t} + \beta_4 \cdot MM_{i,t} + B^{'} \cdot D_i^{ETF} + + C^{'} \cdot D_y^{year} + \varepsilon_{i,t} \tag{4}$$

Return differences $R_{i,t}^{diffCP}$ of ETF i on day t are again explained by DAX returns, return volatility, turnover, product dummies and yearly time dummies. Coefficient β_4 captures the effect of annual return differences induced by market makers, $MM_{i,t}$, of ETF i at time t.

[20] ETF 2 is arbitrarily selected as the reference. For robustness purposes, different references are applied and results (not illustrated) are qualitatively confirmed.

3.2 NAV and CP regression results

Table 3 illustrates the dependence of annual return differences between the DAX and NAVs, where $R_{i,t}^{diffNAV}$ is the dependent variable. The regression specification includes the sample from January 2010 to December 2018. Since ETF 2 is omitted as the reference, the constant term can be interpreted as ETF 2-specific coefficient. Product-specific results for ETFs 4, 6 and 7 show significant negative differences for all models 1 to 4. Negative coefficients indicate that costs of ETFs 4, 6 and 7 significantly exceed those of ETF 2. A negative coefficient of ETF 4 between -0.112% in model 1 and -0.096% in model 4 indicates that annual ETF 4 costs significantly exceed ETF 2 costs by approximately 0.1%. In contrast to the hypothesis as described in Elton, Gruber and Busse (2004), return differences of the DAX ETF market do not converge for the total sample. In line with table 2 results, for all models 1 to 4, the product-specific order remains stable where ETF 6 costs are the highest. Interaction terms (see following sections) contribute to explaining return differences.

It can be observed that ETF coefficients remain stable throughout different models. ETF 3 coefficients, for example, are identical in models 1 and 2 (value 0.008) and in models 3 and 4 (value 0.011). Indeed, coefficients vary slightly throughout models, however, this is not visible due to rounding. Although results remain qualitatively stable, the variation of coefficients slightly increases with the number of AR terms (not illustrated).

Table 3: Dependence of annual return differences between the DAX and NAVs from January 2010 to December 2018

	Model 1 coeff.° (%) (t Stat.)		Model 2 coeff.° (%) (t Stat.)		Model 3 coeff.° (%) (t Stat.)		Model 4 coeff.° (%) (t Stat.)	
Constant	0.062		-0.651	***	-0.289	***	0.227	*
	(1.264)		(-6.203)		(-4.346)		(1.942)	
ETF1	-0.042		-0.041		-0.081		-0.081	
	(-1.005)		(-0.936)		(-0.855)		(-0.916)	
ETF3	0.008		0.008		0.011		0.011	
	(0.187)		(0.176)		(0.240)		(0.259)	
ETF4	-0.112	***	-0.112	**	-0.096	*	-0.096	*
	(-2.675)		(-2.534)		(-1.716)		(-1.850)	
ETF5	0.016		0.016		0.031		0.030	
	(0.383)		(0.362)		(0.577)		(0.605)	
ETF6	-0.410	***	-0.410	***	-0.392	***	-0.393	***
	(-8.167)		(-7.740)		(-5.942)		(-6.386)	
ETF 7	-0.206	***	-0.229	***	-0.220	**	-0.189	**
	(-2.938)		(-3.093)		(-2.551)		(-2.354)	
R_t^{DAX}	-1.374	***					-1.412	***
	(-20.184)						(-19.860)	
σ_t			1.750	***			-0.805	*
			(4.129)				(-1.899)	
$Turn_{i,t}$					0.113		0.110	
					(0.477)		(0.501)	
D_y^{year}	yes		yes		yes		yes	
AR term	yes		yes		yes		yes	
Obs.	13,192		13,192		13,192		13,192	
Adj. R^2	0.799		0.793		0.793		0.799	

° Significance levels are at 10% = *, 5% = ** and 1% = ***.

For DAX index returns, the observed effect is unequivocal. In both models 1 and 4, a highly significant negative coefficient can be observed for R_t^{DAX}. This implies that costs increase with increasing DAX returns. This contradicts the observation by Gil-Bazo and Ruiz-Verdú (2009), who observe a negative dependence between the performance and the fees of mutual funds. An explanation of this effect can be found in cash holdings of ETFs. Since cash does not reflect the development of the underlying index, a reverse relation between costs and the development of the index can be observed. In the case of increasing DAX returns, for example, issuers' management activities cannot fully compensate the effect of cash holdings through income sources such as securities lending. A coefficient of 100% can be interpreted to mean that a one percent increase in DAX returns leads to a one percent increase in negative return differences. However, return differences are moderate since coefficients are -1.374% and -1.412%. This indicates that a one percent increase in DAX returns leads to a 0.014% increase in ETF costs. The results of DAX return volatility as illustrated in models 2 and 4 are not unequivocal. Turnover coefficients of models 3 and 4 are positive but insignificant. In contrast to LaPlante (2001) or Rompotis (2012), economies of scale that lead to reduced costs with increases in fund size are not observable.

Since ETFs cannot be traded at fair value for private investors, return differences are analysed using CPs. The results can be interpreted as a robustness check of the results in table 3. Furthermore, regression (4) specifications make it possible to identify the influence of market makers on return differences. Therefore, deviations between $R_{i,t}^{diffCP}$ and $R_{i,t}^{diffNAV}$, denominated as $MM_{i,t}$ are calculated for each ETF, where $R_{i,t}^{diffCP}$ is deducted from $R_{i,t}^{diffNAV}$. Hence, a positive value indicates that costs relying on CPs are higher compared to costs relying on NAVs and vice versa. Table 4 briefly illustrates descriptive statistics of $MM_{i,t}$. Mean values range between -0.011% for ETF 3 and 0.018% for ETF 7. On average, CP returns are closely related to NAV returns. Comparable results can be observed for medians. However, with a range between -3.419% as a minimum and 2.894% as a maximum, and for standard deviations up to 0.339%, $MM_{i,t}$ outcomes indicate that market makers can considerably influence return differences.

Table 4: Market-maker-induced deviations (%) of annual CP returns from NAV returns from January 2010 to December 2018

	MM1	**MM2**	**MM3**	**MM4**	**MM5**	**MM6**	**MM7**
Min.	-1.662	-0.802	-2.852	-2.477	-3.419	-0.409	-1.026
Max.	0.625	0.886	2.577	2.223	2.894	0.815	2.873
Mean	-0.002	-0.003	-0.011	-0.004	0.003	0.003	0.018
Median	0.000	-0.003	-0.004	-0.003	-0.002	0.000	0.000
Std.dev.	0.106	0.124	0.221	0.166	0.339	0.099	0.232
Observations	2,281	2,281	2,281	2,281	2,277	1,267	524

Table 5 illustrates the dependence of $R_{i,t}^{diffCP}$ on product specifications, annual DAX returns, DAX return volatility, relative turnover and market-maker-induced deviations from NAV returns. Results show that product effects qualitatively confirm the results in table 3. Significant return differences relying on CPs can be observed between ETF products. In line with the results in tables 2 and 3, the product-specific order remains stable. Again, return differences of ETFs 4, 6 and 7 significantly exceed those of ETF 2 as the reference.

Again, the coefficient of R_t^{DAX} is highly significant and negative (see models 2 and 5). Coefficient results of σ_t are not unequivocal. In both models 3 and 5, $Turn_{i,t}$ coefficient results are not significant, confirming the outcomes in table 3.

In models 1 and 5, $MM_{i,t}$ coefficients are positive and highly significant. A positive coefficient can be interpreted that market-makers contribute to increase costs and income. Coefficients of models 1 and 5 are 53.206% and 52.960% which means that a one percent increase in $MM_{i,t}$ leads to a 0.53206% and 0.52960% increase in return differences. Product-specific market-maker effects are illustrated in more detail in section 3.4.

Table 5: Dependence of annual return differences between the DAX and CPs from January 2010 to December 2018

	Model 1 coeff.° (%) (t Statistic)		Model 2 coeff.° (%) (t Statistic)		Model 3 coeff.° (%) (t Statistic)		Model 4 coeff.° (%) (t Statistic)		Model 5 coeff.° (%) (t Statistic)	
Constant	-0.229	***	0.117	**	-0.571	***	-0.266	***	0.311	**
	(-3.757)		(2.368)		(-5.556)		(-4.172)		(2.390)	
ETF1	-0.038		-0.039		-0.038		-0.078		-0.098	
	(-0.666)		(-0.938)		(-0.904)		(-0.857)		(-0.847)	
ETF3	0.014		0.018		0.018		0.021		0.018	
	(0.242)		(0.441)		(0.434)		(0.497)		(0.333)	
ETF4	-0.105	*	-0.107	**	-0.107	**	-0.091	*	-0.083	
	(-1.836)		(-2.570)		(-2.527)		(-1.702)		(-1.233)	
ETF5	0.025		0.017		0.017		0.032		0.042	
	(0.430)		(0.405)		(0.403)		(0.620)		(0.648)	
ETF 6	-0.379	***	-0.421	***	-0.423	***	-0.406	***	-0.344	***
	(-5.559)		(-8.428)		(-8.328)		(-6.420)		(-4.332)	
ETF 7	-0.382	***	-0.225	***	-0.244	***	-0.232	***	-0.289	***
	(-4.051)		(-3.214)		(-3.433)		(-2.814)		(-2.812)	
$MM_{i,t}$	53.206	***							52.960	***
	(109.902)								(112.450)	
R_t^{DAX}			-1.502	***					-1.635	***
			(-21.269)						(-27.540)	
σ_t					1.487	***			-0.868	*
					(3.568)				(-1.921)	
$Turn_{i,t}$							0.112		0.159	
							(0.496)		(0.553)	
D_y^{year}	yes		yes		yes		yes		yes	
AR terms	yes		yes		yes		yes		yes	
Obs.	13,192		13,192		13,192		13,192		13,192	
Adj. R^2	0.879		0.778		0.771		0.771		0.886	

° Significance levels are at 10% = *, 5% = ** and 1% = ***.

3.3 Interaction terms

To illustrate the product-specific variable impact of annual DAX returns and market makers, interaction terms are applied relying on regression specifications (3) and (4). In order to analyse the variable effect of DAX returns, equation (3) is extended as follows:

$$R_{i,t}^{diffNAV} = \beta_0 + \beta_1 \cdot R_t^{DAX} + E' \cdot R_t^{DAX} \cdot D_i^{ETF} + \beta_2 \cdot \sigma_t + \beta_3 \cdot Turn_{i,t} + $$
$$+ B' \cdot D_i^{ETF} + C' \cdot D_y^{year} + \varepsilon_{i,t} \quad (5)$$

E is an $(I-1) \times 1$ vector of coefficients that captures the product-specific variable impact of DAX returns, R_t^{DAX}, and ETFs i. D_i^{ETF} is the corresponding $(I-1) \times 1$ vector of product dummies.

In the following, the product-specific variable impact of market-maker-induced costs is analysed:

$$R_{i,t}^{diffCP} = \beta_0 + \beta_1 \cdot R_t^{DAX} + \beta_2 \cdot \sigma_t + \beta_3 \cdot Turn_{i,t} + \beta_4 \cdot MM_{i,t} + F' \cdot MM_{i,t} \cdot D_i^{ETF} + $$
$$+ B' \cdot D_i^{ETF} + C' \cdot D_y^{year} + \varepsilon_{i,t} \quad (6)$$

F is an $(I-1) \times 1$ vector of coefficients that captures any product-specific variable impact of market-maker-induced costs, $MM_{i,t}$. D_i^{ETF} is the corresponding $(I-1) \times 1$ vector of product dummies.

Table 6 illustrates the dependence of annual return differences between NAV and CP returns versus DAX index returns. Model 1 results are based on equation (5), where $R_{i,t}^{diffNAV}$ is the dependent variable. Model 2 results are based on equation (6), where $R_{i,t}^{diffCP}$ is the dependent variable.

Model 1 captures any product-specific variable impact of annual DAX returns. ETF-specific fixed effects range between -0.100% and 0.138%. Since ETF 2 is omitted as the reference, variable product effects are interpreted relative to the coefficient R_t^{DAX}. For R_t^{DAX}, again a highly significant negative impact can be observed. The variable impact of ETF 4 does not significantly differ from the reference. For ETFs 1, 3, 5, 6 and 7, interaction terms are highly significant and negative. Hence, negative return differences of ETFs 1, 3, 5, 6 and 7 increase more strongly with increasing DAX returns than those of ETFs 2 and 4. It is interesting to observe that interaction coefficients of ETFs 6 (-6.225%) and 7 (-2.458%) considerably exceed those of other ETFs. An explanation can be found in the use of profits, since ETFs 6 and 7 are distributing. For distributing funds, dividends are paid out to investors and not reinvested in the underlying benchmark. To make funds comparable, dividends are added to NAVs and CPs. This corresponds to cash holdings and return differences of distributing ETFs are more sensitive to returns of the underlying benchmark than accumulating ETFs. Although ETF 5 is distributing, the interaction coefficient is qualitatively low. An explanation can be found in a comparatively low number of dividend payments.

Model 2 captures any product-specific variable impact of market-maker-induced return differences. In line with the results of table 5, a highly significant positive impact of $MM_{i,t}$ can be observed. Variable product effects are interpreted relative to the coefficient $MM_{i,t}$ which relates to ETF 2 as the reference. A highly significant and positive value of 80.426% shows that market makers of ETF 2 considerably influence return differences. Interaction terms of ETFs range between -72.333% and 19.496%. Results show that market makers significantly contribute to return differences, however, to varying degrees.

Table 6: Dependence of annual return differences between NAV and CP returns versus DAX index returns from January 2010 to December 2018

	Model 1 coeff.° (%) (t Statistic)		Model 2 coeff.° (%) (t Statistic)	
Dependent variable	$R_{i,t}^{diffNAV}$		$R_{i,t}^{diffCP}$	
Interaction term	$R_{DAX,t}$		$MM_{i,t}$	
Constant	0.024		0.070	
	(0.374)		(0.484)	
ETF1	0.044		-0.150	
	(1.008)		(-0.998)	
ETF3	0.064	***	0.005	
	(2.722)		(0.074)	
ETF4	-0.100	***	-0.076	
	(-3.618)		(-0.879)	
ETF5	0.064	**	0.072	
	(2.402)		(0.871)	
ETF6	0.138	***	-0.268	***
	(4.335)		(-2.684)	
ETF7	0.085	*	-0.500	***
	(1.954)		(-3.877)	
$\boldsymbol{MM_{i.t}}$			80.462	***
			(53.175)	
$\boldsymbol{R_t^{DAX}}$	-0.443	***	-1.348	***
	(-4.997)		(-28.100)	
ETF1 x interaction term	-0.653	***	6.497	***
	(-5.618)		(2.846)	
ETF3 x interaction term	-0.484	***	19.496	***
	(-4.166)		(11.482)	
ETF4 x interaction term	-0.133		-4.485	**
	(-1.145)		(-2.425)	
ETF5 x interaction term	-0.442	***	-55.869	***
	(-3.798)		(-34.898)	
ETF6 x interaction term	-6.225	***	-27.420	***
	(-41.310)		(-9.700)	
ETF7 x interaction term	-2.458	***	-72.333	***
	(-11.272)		(-33.169)	
$\boldsymbol{\sigma_t}$	-0.639	***	-0.262	
	(-2.787)		(-0.563)	
$\boldsymbol{Turn_{i.t}}$	-0.025		0.261	
	(-0.242)		(0.710)	
$\boldsymbol{D_y^{year}}$	yes		yes	
AR-terms	yes		yes	
Obs.	13,192		13,192	
Adj. R^2	0.816		0.931	

° Significance levels are at 10% = *, 5% = ** and 1% = ***.

4. Robustness

In this section, the robustness of the results is examined. Robustness tests illustrated as follows include subsample periods and regression specifications with dummy variables for DAX index returns. Further, regression specifications (not illustrated) are estimated with different AR terms and Newey and West (1987) is employed to correct for heteroscedasticity and autocorrelations in the sample's residuals. Robustness regressions qualitatively confirm the results.

4.1 Subsample periods

The total sample includes nine years of annual return differences between January 2010 and December 2018. In order to analyse whether results are robust over time, the total sample is divided into two subsamples that equal 4.5 years each. Subsample T1 includes the time interval from January 2010 to June 2014, and subsample T2 the time interval from July 2014 to December 2018. T1 includes 5,847 observations and T2 7,345 observations. T1 and T2 add up to 13,192 observations as included in the regressions in tables 3, 5 and 6. T2 observations exceed those of T1 since ETFs 6 and 7 are issued in 2012 and 2015 respectively, and return differences are available since 2013 and 2016. Regression specifications correspond to equations (3) and (4), and are estimated for $R_{i,t}^{diffCP}$ and $R_{i,t}^{diffNAV}$ as dependent variables. Table 7 shows the results.

Table 7: Dependence of annual return differences between the DAX and NAVs, and the DAX and CPs during time intervals T1 from January 2010 to June 2014, and T2 from July 2014 to December 2018

	Model 1 coeff.° (%) (t Statistic)		**Model 2** coeff.° (%) (t Statistic)		**Model 3** coeff.° in % (t Statistic)		**Model 4** coeff.° (%) (t Statistic)	
	$R_{i,t}^{diffNAV}$		$R_{i,t}^{diffNAV}$		$R_{i,t}^{diffCP}$		$R_{i,t}^{diffCP}$	
	T1		**T2**		**T1**		**T2**	
Constant	-0.104	**	0.093	***	-0.098	**	-0.258	
	(-2.275)		(0.533)		(-2.240)		(-0.782)	
ETF1	-0.088	**	-0.081		-0.091	***	-0.138	
	(-2.451)		(-0.410)		(-2.628)		(-0.250)	
ETF3	0.054	***	-0.043		0.060	***	-0.061	
	(3.310)		(-0.453)		(3.810)		(-0.209)	
ETF4	-0.120	***	-0.073		-0.118	***	-0.055	
	(-5.785)		(-0.678)		(-5.881)		(-0.172)	
ETF5	0.119	***	-0.055		0.117	***	-0.007	
	(6.111)		(-0.525)		(6.268)		(-0.021)	
ETF6	-0.304	***	-0.381	***	-0.313	***	0.061	
	(-8.482)		(-3.517)		(-9.051)		(0.197)	
ETF7			-0.258	***			-1.094	***
			(-1.943)				(-3.177)	
$\boldsymbol{MM_{i,t}}$					66.429	***	23.930	***
					(96.631)		(41.686)	
$\boldsymbol{R_t^{DAX}}$	-0.565	***	-2.427	***	-0.560	***	-2.217	***
	(-17.164)		(-20.781)		(-17.824)		(-32.729)	
$\boldsymbol{\sigma_t}$	-0.258	*	-0.483	***	-0.292	**	0.219	
	(-1.733)		(-0.595)		(-2.037)		(0.221)	
$\boldsymbol{Turn_{i,t}}$	0.035		0.192		0.043		0.336	
	(0.326)		(0.442)		(0.408)		(0.287)	
$\boldsymbol{D_y^{year}}$	yes		yes		yes		yes	
AR terms	yes		yes		yes		yes	
Obs.	5,847		7,345		5,847		7,345	
Adj. R²	0.442		0.877		0.710		0.961	

° Significance levels are at 10% = *, 5% = ** and 1% = ***.

One can observe that product-specific ETF coefficients significantly differ during T1. During T2, significance levels largely decrease, which can be interpreted to mean that product differences tend to converge over time. Results are valid for both $R_{i,t}^{diffCP}$ and $R_{i,t}^{diffNAV}$ as dependent variables. For ETF 6 (with $R_{i,t}^{diffNAV}$ as dependent variable) and ETF 7, significant differences can be observed also in T2. The construction as distributing fund delivers an argument for these differences. As Figure 1 illustrates, return differences vary highly for ETF 6 as distributing fund. The same is valid for ETF 7. Again, the coefficient R_t^{DAX} is highly significant and negative during T1 and T2. Market makers significantly contribute to return differences, however, the coefficient decreases in T2.

Table 8 illustrates regression specifications of equations (5) and (6) including interaction terms. Results largely confirm table 7 coefficients. During T2, ETF coefficients tend to converge. The coefficient of R_t^{DAX} is negative and significant during T1 and T2. This is valid for both $R_{i,t}^{diffCP}$ and $R_{i,t}^{diffNAV}$ as dependent variables. Model 2 shows that in T2 the variable impact of R_t^{DAX} is highest for distributing funds. This is not the case during T1 (see model 1). An argument delivers the late issuance date of ETFs 6 and 7 and a low number of dividend payments for distributing funds during T1. Again, market makers significantly contribute to increasing return differences, however, to varying degrees (see models 3 and 4). The variable impact of ETF 6 is positive but close to zero during T2 and does not significantly differ from ETF 2. Overall, the influence of market makers is highly significant in T1 and T2. A slight decrease in coefficients can be observed in T2, for example, coefficient $MM_{i,t}$ decreases from 90.700% in model 3 to 57.661% in model 4.

Table 8: Dependence of annual return differences between NAV and CP returns versus DAX index returns from January 2010 to December 2018

	Model 1 coeff.° (%) (t Statistic)		Model 2 coeff.° (%) (t Statistic)		Model 3 coeff.° (%) (t Statistic)		Model 4 coeff.° (%) (t Statistic)	
Dependent variable	$R_{i,t}^{diffNAV}$		$R_{i,t}^{diffNAV}$		$R_{i,t}^{diffCP}$		$R_{i,t}^{diffCP}$	
Interaction term	$R_{DAX,t}$		$R_{DAX,t}$		$MM_{i,t}$		$MM_{i,t}$	
	T1		T2		T1		T2	
Constant	-0.179	***	0.046		-0.106	**	-0.409	
	(-4.505)		(0.663)		(-2.202)		(-1.068)	
ETF1	0.006		0.006		-0.102	***	-0.383	
	(0.174)		(0.100)		(-2.620)		(-0.572)	
ETF3	0.188	***	-0.014		0.058	***	-0.096	
	(10.281)		(-0.423)		(3.283)		(-0.247)	
ETF4	-0.100	***	-0.070	*	-0.113	***	0.019	
	(-4.650)		(-1.882)		(-5.012)		(0.045)	
ETF5	0.146	***	0.004		0.121	***	0.097	
	(7.267)		(0.099)		(5.720)		(0.236)	
ETF6	-0.188		0.070	*	-0.317	***	0.246	
	(-1.596)		(1.875)		(-8.178)		(0.617)	
ETF7			0.060				-1.136	***
			(1.213)				(-2.645)	
$MM_{i,t}$					90.700	***	57.661	***
					(42.908)		(28.894)	
R_t^{DAX}	-0.208	***	-0.684	***	-0.560	***	-2.009	***
	(-4.130)		(-4.505)		(-16.795)		(-34.454)	
ETF1 x interaction term	-0.749	***	-0.379	*	3.721		18.512	***
	(-11.439)		(-1.939)		(1.125)		(6.472)	
ETF3 x interaction term	-0.818	***	-0.271		9.832	***	39.031	***
	(-12.302)		(-1.387)		(4.267)		(12.104)	
ETF4 x interaction term	-0.059		-0.279		8.624	***	-29.245	***
	(-0.890)		(-1.427)		(3.332)		(-11.913)	
ETF5 x interaction term	-0.115	*	-1.004	***	-58.384	***	-48.727	***
	(-1.745)		(-5.119)		(-26.097)		(-23.058)	
ETF6 x interaction term	-0.569		-6.901	***	-48.064	***	0.000	
	(-1.065)		(-35.334)		(-8.638)		(0.000)	
ETF7 x interaction term			-2.580	***			-48.944	***
			(-10.217)				(-21.066)	
σ_t	-0.239	*	-0.884	**	-0.312	**	0.446	
	(-1.889)		(-2.415)		(-1.986)		(0.472)	
$Turn_{i,t}$	0.108		0.049		0.080		0.911	
	(1.150)		(0.357)		(0.678)		(0.675)	
D_y^{year}	yes		yes		yes		yes	
AR-terms	yes		yes		yes		yes	
Obs.	5,847		7,345		5,847		7,345	
Adj. R^2	0.465		0.891		0.828		0.971	

° Significance levels are at 10% = *, 5% = ** and 1% = ***.

Product-specific return differences largely decrease during T2. According to the hypothesis of Elton, Gruber and Busse (2004), when prices of identical products are expected to converge for highly liquid and transparent markets, results can be interpreted such that DAX ETF markets tend to become more efficient over time. However, the influence of DAX index returns and market makers on return differences remains significant.

4.2 Index return dummies

Dependent variables $R_{i,t}^{diffCP}$ and $R_{i,t}^{diffNAV}$ are constructed by subtracting annual DAX index returns from annual NAV and CP returns. This allows costs to be accounted as negative deviation, and income as positive deviation. Relying on Gil-Bazo and Ruiz-Verdú (2009) and to account for cash holdings, DAX index returns are included as an independent variable. Although DAX returns are subtracted from NAV and CP returns as dependent variable to avoid endogeneity, robustness tests are applied to confirm results. In order to show the econometric robustness of regression results, R_t^{DAX} is replaced by a dummy variable, denoted as $D_{up}^{R_{DAX}}$. The dummy variable is constructed to be one in the case of positive annual DAX returns and zero elsewhere. Due to cash holdings of ETF products, $D_{up}^{R_{DAX}}$ coefficients are expected to be negative. Robustness tests include NAV and CP results as well as different time periods. Table 9 corresponds to equation (3) and shows the results of $R_{i,t}^{diffNAV}$ as a dependent variable for the total sample and during T1 and T2. For all models, $D_{up}^{R_{DAX}}$ coefficients are negative and highly significant. Again, ETF-specific coefficients converge over time.

Table 9: Dependence of annual return differences between the DAX and NAVs during time intervals from January 2010 to December 2018 (Total), from January 2010 to June 2014 (T1), and from July 2014 to December 2018 (T2)

	Model 1 coeff.° (%) (t Statistic)		**Model 2** coeff.° (%) (t Statistic)		**Model 3** coeff.° (%) (t Statistic)	
	Total		**T1**		**T2**	
Constant	-0.483	***	-0.131	**	-0.687	***
	(-4.253)		(-2.213)		(-3.766)	
ETF1	-0.081		-0.088	**	-0.086	
	(-0.883)		(-2.255)		(-0.400)	
ETF3	0.011		0.054	***	-0.043	
	(0.250)		(3.037)		(-0.423)	
ETF4	-0.096	*	-0.120	***	-0.071	
	(-1.793)		(-5.286)		(-0.608)	
ETF5	0.030		0.119	***	-0.053	
	(0.584)		(5.607)		(-0.467)	
ETF6	-0.393	***	-0.305	***	-0.367	***
	(-6.198)		(-7.794)		(-3.112)	
ETF7	-0.203	**			-0.308	**
	(-2.449)				(-2.129)	
$D_{up}^{R_{DAX}}$	-0.089	***	-0.152	***	-0.085	***
	(-7.775)		(-9.363)		(-7.281)	
σ_t	1.287	***	-0.071		3.169	***
	(3.075)		(-0.384)		(3.775)	
$Turn_{i,t}$	0.111		0.037		0.205	
	(0.488)		(0.310)		(0.433)	
D_y^{year}	yes		yes		yes	
AR terms	yes		yes		yes	
Obs.	13,192		5,847		7,345	
Adj. R^2	0.794		0.425		0.871	

° Significance levels are at 10% = *, 5% = ** and 1% = ***.

As illustrated in table 10, the same is valid for $R_{i,t}^{diffCP}$ as a dependent variable which corresponds to equation (4). $D_{up}^{R_{DAX}}$ coefficients are again negative and highly significant during the total sample and during subsamples. Again, product-specific deviations significantly decrease from T1 to T2. A highly significant influence of market makers on return differences can be observed. In line with preceding results, market-maker coefficients decrease over time.

Table 10: Dependence of annual return differences between the DAX and CPs during time intervals from January 2010 to December 2018 (Total), from January 2010 to June 2014 (T1), and from July 2014 to December 2018 (T2)

	Model 1 coeff.° (%) (t Statistic)		Model 2 coeff.° (%) (t Statistic)		Model 3 coeff.° (%) (t Statistic)	
	Total		**T1**		**T2**	
Constant	-0.549	***	-0.150	***	-1.126	***
	(-4.282)		(-2.655)		(-2.693)	
ETF1	-0.093		-0.091	**	-0.130	
	(-0.798)		(-2.429)		(-0.179)	
ETF3	0.018		0.060	***	-0.063	
	(0.325)		(3.511)		(-0.158)	
ETF4	-0.084		-0.117	***	-0.058	
	(-1.231)		(-5.398)		(-0.133)	
ETF5	0.041		0.118	***	0.015	
	(0.621)		(5.775)		(0.035)	
ETF6	-0.353	***	-0.314	***	0.231	
	(-4.397)		(-8.350)		(0.552)	
ETF7	-0.319	***			-1.431	***
	(-3.058)				(-3.120)	
$\boldsymbol{MM_{i,t}}$	53.285	***	66.428	***	23.498	***
	(110.169)		(95.673)		(38.378)	
$\boldsymbol{D_{up}^{R_{DAX}}}$	-0.062	***	-0.142	***	-0.044	***
	(-7.067)		(-9.275)		(-6.629)	
$\boldsymbol{\sigma_t}$	1.579	***	-0.038		4.104	***
	(3.516)		(-0.216)		(3.646)	
$\boldsymbol{Turn_{i,t}}$	0.154		0.044		0.314	
	(0.532)		(0.385)		(0.209)	
$\boldsymbol{D_y^{year}}$	yes		yes		yes	
AR terms	yes		yes		yes	
Obs.	13,192		5,847		7,345	
Adj. R^2	0.880		0.700		0.956	

° Significance levels are at 10% = *, 5% = ** and 1% = ***.

5. Conclusions

In recent years, ETFs have become a popular instrument for professional and private investors. Financial literature addresses topics such as tracking quality as well as arbitrage and hedging opportunities, which are primarily relevant for professional investors. While the focus of these studies is short-term, the present paper focuses on the perspective of long-term investors.

Prominent studies, such as Elton, Gruber, Comer and Lee (2002), Lang and Röder (2008), and Blitz, Huij and Swinkels (2012), address ETF-related costs or income relying on annual financial statements. However, this paper is the first to investigate a daily development of annual return differences between DAX ETFs and the DAX index. To address the perspective of long-term investors and in line with the purpose of passively managed products, annual ETF returns are compared with corresponding index returns as a benchmark. This allows costs to be accounted as negative deviations between ETF returns and index returns, and income as positive deviations. According to Baltussen, Post and Vliet (2012) and in line with financial reporting standards, annual periods are selected to cover total costs and income. The construction of a daily rolling window of annual return differences contributes (1) to identify long-term costs and income beyond annual financial statements (2) to identify long-term costs and income on a daily level and (3) to apply statistical tests to capture product specifications.

The seven analysed ETFs meet high liquidity and transparency standards and the sample covers a broad market spectrum that includes all market leading DAX ETFs as well as products with a small market share. Based on the hypothesis of Elton, Gruber and Busse (2004), prices of identical products are expected to converge for highly liquid and transparent markets. The present study analyses annual return differences between DAX ETFs and the DAX index to identify whether costs of DAX ETFs converge or diverge. Additionally, the influence of DAX index returns, index return volatility, relative annual turnover and market makers is illustrated.

On average, NAV-based annual return differences range between -0.234% and -0.699%. Average deviations based on CPs range between -0.230% and -0.701%. Although return differences are qualitatively low, they considerably exceed TERs, which range between 0.08% and 0.16% p.a. Regressions reveal that return differences significantly differ between products for the total sample between January 2010 and December 2018. The same holds for subsample T1 between January 2010 and June 2014. However, T2 results between July 2014 and December 2018 show that ETF-specific return differences tend to converge over time. This is in line with the hypothesis of Elton, Gruber and Busse (2004), who expect prices of identical products to converge for highly liquid and transparent markets. Deviations for ETFs 6 and 7 can be explained by fund characteristics such as dividend distribution.

Regression results show that DAX index returns significantly influence return differences between DAX ETFs and the DAX index. In contrast to Gil-Bazo and Ruiz-Verdú (2009), who observe a negative dependence between the performance and the fees of mutual funds, DAX ETF costs significantly increase when index returns increase. Cash holdings deliver a plausible argument since cash does not increase with index returns. The effect of DAX index returns is significantly negative for all ETFs. One argument is that, issuers' management activities, e.g., securities lending do not fully compensate negative return differences due to cash holdings. Interaction terms between ETFs and annual DAX returns reveal that the influence of DAX index returns strongly varies between ETFs. An argument can be found in the use of profits between distributing or accumulating funds. In contrast to accumulating funds, dividend payments of distributing funds are not reinvested in the index. Hence, dividend payments can be interpreted as cash holdings since they are paid out to investors. To account for this, NAVs and CPs are adjusted for dividend payments which contributes to an increasing sensitivity for DAX index returns.

Return differences are analysed for NAVs and CPs. Since CP results largely correspond to NAV results, long-term costs of institutional investors and private investors are comparable. For market makers, a significant influence on return differences can be observed. An analysis of time intervals T1 and T2 reveals that the influence of market-makers slightly decrease over time. Also products tend to converge over time which indicates that the DAX index market becomes more efficient. However, return differences of ETFs are significantly influenced by DAX index returns. This is valid for the total sample and during subsample periods. Besides cash holdings, fund characteristics such as dividend distribution contributes to return differences. Hence, ETF characteristics can be relevant for investment decisions especially in the context of future market expectations.

References

Alexander C, Barbosa A (2008) Hedging index exchange traded funds. J Bank Financ 32:326–337

Anderson NL, Highfield MJ (2014) ETF Expense Ratios: Are All Costs Borne by the Investor? An Empirical Analysis Using the Open-End Mutual Fund Model. Available via DIALOG:. Accessed 9 Jan 2019

Baltussen G, Post T, Vliet P (2012) Downside Risk Aversion, Fixed Income Exposure, and the Value Premium Puzzle. J Bank Finance 36(12): 3382–3398

Ben-David F, Franzoni F, Moussawi R (2017) Exchange Traded Funds. Annual Rev Financ Econ 9: 169–189

Blitz D, Huij J, Swinkels L (2012) The Performance of European Index Funds and Exchange Traded Funds. Eur Financ Manage 18(4): 649–662

Breusch T (1978) Testing for autocorrelation in dynamic linear models. Aust Econ Pap 17:334–355

Chen SY, Lin CC, Cou PH, Hwang DY (2002) A comparison of hedge effectiveness and price discovery between TAIFEX TAIEX index futures and SGX MSCI Taiwan index futures. Rev Pac Basin Financ Mark Policies 5:277–300

Cherry J (2004) The limits of arbitrage: evidence from exchange traded funds. Available via DIALOG: https://ssrn.com/abstract=628061. Accessed 9 Jan 2019

Cremers M, Petajisto A (2009) How Active is Your Fund Manager? A New Measure That Predicts Performance. Rev Financ Stud 22(9): 3329–3365

Deutsche Bundesbank (2015) Das Spar- und Anlageverhalten privater Haushalte in Deutschland vor dem Hintergrund des Niedrigzinsumfelds. Technical Report: 13 - 32

Dorfleitner G, Gerl A, Gerer J (2018) The Pricing Efficiency of Exchange-Traded Commodities. Rev Manag Sci 12:255–284

Elton EJ, Gruber MJ, Busse JA (2004) Are Investors Rational? Choices among Index Funds. J Finance 59(1): 261–288

Elton EJ, Gruber MJ, Comer G, Li K (2002) Spiders: Where are the Bugs? J Business 75:453–472

Engle R, Sarkar D (2006) Premiums-discounts and exchange traded funds. J Derivatives 13:27–45

French KR, Schwert W, Stambaugh RF (1987) Expected stock returns and volatility. J Financ Econ 19: 3–29

Ghysels E, Santa-Clara P, Valkanov R (2005) There is a Risk-Return trade-off After All. J Financ Econ 76(3): 509–548

Gil-Bazo J, Ruiz-Verdú P (2009) The Relation between Price and Performance in the Mutual Fund Industry. J Finance 64(5): 2153–2183

Godfrey L (1978) Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. Econometrica 46:1293–1301

Hasbrouck J (2003) Intraday price formation in U.S. equity index markets. J Finance 58:2375–2399

Johnson WF (2008) Tracking Errors of Exchange Traded Funds. J Asset Manage 10:253–262

Klein C, Kundisch D (2009) Der Tracking Error von indexnachbildenden Instrumenten auf den DAXeine empirische Analyse des DAX EX sowie zehn Indexzertifikaten. Der Betrieb 62:1141–1145

Lang SE, Röder K (2008) Die Kosten des Indextrackings: eine Fallstudie über den Exchange Traded Fund DAX®EX. Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung 60(3): 298–321

LaPlante (2001) Influences and Trends in Mutual Fund Expense Ratios. J Financ Res 24(1): 45–63

Latzko DA (1999) Economies of Scale in Mutual Fund Administration. J Financ Res 22(3): 331–339

Malhotra DK, McLeod RW (1997) An Empirical Analysis of Mutual Fund Expenses. J Financ Res 20(2): 175–190

Markowitz H (1952) Portfolio Selection. J Finance 7: 77 – 91

Merton RC (1980) On Estimating the Expected Return on the Market. J Financ Econ 323–361

Merz T (2015) Tracking Risk of Exchange Traded Funds Revisited - A Multivariate Regression Approach. Available via DIALOG: https://ssrn.com/abstract=2553766. Accessed 9 Jan 2019

Newey W, West K (1987) A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55:703–708

Osterhoff F, Kaserer C (2016). Determinants of Tracking Error in German ETFs - the Role of Market Liquidity, Managerial Finance 42: 417–437

Petajisto A (2017) Inefficiencies in the Pricing of Exchange-Traded Funds. Financ Analyst J 73(1): 24–54

Richie N, Daigler R, Gleason K (2008) The limits to stock index arbitrage: examining S&P 500 futures and SPDRS. J Futures Mark 28:1182–1205

Roll R (1992) A Mean/Variance Analysis of Tracking Error. J Portf Manage 18(4): 13–22

Rompotis GG (2006) A Empirical Look on Exchange Traded Funds. Available via

DIALOG: https://ssrn.com/abstract=905770. Accessed 9 Jan 2019

Rompotis GG (2012) The German Exchange Traded Funds. J Appl Finance 18(4): 62–82

Schmidhammer C, Lobe S, Röder K (2014) The real benchmark of DAX index products and the influence of information dissemination: A natural experiment, J Asset Manage 15(2): 129–149

Schmidhammer C (2018) Performance of Bond Ladder Strategies: Evidence from the German Stock Market. Credit and Capital Markets 51(3): 421–443

Sharpe WF (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. J Financ 19(3): 425–442

Tobin J (1958) Liquidity Preference as Behavior Towards Risk. Rev Econ Stud 25(2): 65–86