EUROPEAN BOND ETFs - TRACKING ERRORS AND SOVEREIGN DEBT CRISIS

Mikica Drenovak,* Branko Urošević, and Ranko Jelic*****

ABSTRACT

We examine effects of the new risk and return paradigm in the Euro sovereign bond market on the tracking performance of 31 Euro zone sovereign debt exchange traded index funds (ETFs), during 2007-2010. The tracking performance was examined using traditional, OLS, and cointegration tracking error models. Overall, ETFs underperform their respective benchmarks. There are, however, some important differences across families of sample ETFs. In particular, ETFs with the highest tracking errors estimated using short term correlations tend to have lowest tracking errors based on cointegration metric. The results of our panel data analysis document significant changes in the sample ETFs' tracking performance during sovereign debt crisis. Our results also confirm that, as a result of the crisis, credit risk considerations have become an important determinant of the ETFs' tracking performance. We also find evidence for the importance of ETFs' replication methods and volatility of underlying indices for the tracking performance, irrespective of the error metric. In an environment of widening sovereign credit default swap (CDS) spreads and divergent yield trends, understanding the credit quality of various issuers together with the selection rules of benchmark indices is, therefore, crucial for understanding ETFs' performance.

KEY WORDS: exchange-traded funds, tracking errors, fixed-income, sovereign debt
JEL CLASSIFICATION: D81, E43, G15, G24

Acknowledgements: The authors are indebted to Drago Indjic and Miles Kumaresan for their valuable suggestions and generous help with data collection. We also thank Carsten Sprenger, Felix Goltz, Chien-Hui Liao, Florian Bitsch, and participants of European Financial Management Symposium on Alternative Investments, Toronto, 2011, for valuable comments. Mikica Drenovak and Branko Urošević gratefully acknowledge the support of the Serbian Ministry of Science and Technology, Grant OH 179005.

*Faculty of Economics, University of Kragujevac, Serbia; mdrenovak@kg.ac.rs.
** Faculty of Economics, University of Belgrade, Serbia, National Bank of Serbia, and the Laboratory for Financial Economics, ICEF, Moscow, Russia; branko.urosevic@cif.co.rs.
***Corresponding author: Ranko Jelic; r.jelic@bham.ac.uk; tel: 44 (0)121 4145990; fax: 44(0)121 4146238; University of Birmingham, Business School, Birmingham, B15 2TT, UK.

1. Introduction

In times of market distress and high uncertainty, investors typically rebalance their portfolios towards less risky and more liquid assets (Longstaff, 2004). The '*flight to quality*' associated with the recent financial crisis spurred the development of the Euro zone sovereign debt exchange traded funds (ETFs) market which today represents one of the fastest growing segments of the overall European ETF market.[1] For example, in 2009 these products gathered \$6.2bn in new cash which was 7.5% of total inflows in all European ETFs. [2] This trend continued in the first half of 2010 with \$5.3bn in new cash direct to this asset class. The sheer number of such funds, their liquidity, and significant total assets under management make them an important new investment class.

The sovereign debt crisis also contributed to the higher volatility and changes in relative importance of risk factors, thus creating a new risk return paradigm in the EU sovereign debt market. Before the crisis risk of Euro sovereign bond indices was low and almost entirely attributable to yield curve risk. After the crisis, the risk of Euro sovereign indices rose by approximately 30% mostly due to the increase in spread levels and volatility (Nomura, 2011). Sovereign bonds from so-called peripheral EU

---

[1] We take the end of August 2008 (i.e. collapse of Lehman Brothers) as the beginning of the financial crisis. Since then, several EU governments have implemented various measures aiming to stabilize their troubled financial systems which was further followed by the divergence in the credit spreads of EU countries. The crisis culminated in early 2010 with the Greek sovereign debt crisis.

[2] By mid 2009, there were 753 registered European exchange traded funds (ETFs) with assets under management (AUM) of over US\$ 183 billion. The proportion of fixed income ETFs in the total ETF assets grew from 5% in 2003 to more than 25% in 2009.Calculations by the authors, based on data from Barclays Global Investors (2009) and Flood (2010b).

countries (i.e. Belgium, Greece, Italy, Ireland, Portugal, and Spain) have become more akin to corporate bonds. The above changes created a new risk return paradigm in the Euro sovereign debt market and call for a more sophisticated approach to assessment of ETFs' performance.

Despite their increasing importance, there is still a paucity of research on the European ETFs and the analysis of ETF tracking errors remains a widely misunderstood and frustrating process for investors (Flood 2010a). Previous studies of the Europe-based ETF tracking errors primarily focused on equity ETFs (Rompotis, 2008; Blitz et al., 2010) and the role of European ETFs in providing investor exposure to various asset classes (Amenc and Goltz, 2009).[3] Recently, Drenovak and Urošević (2011) discussed institutional issues related to the Euro zone sovereign debt ETF market and perform a preliminary ranking of funds based on the standard correlation-based tracking error model, for the period 2007-2009.

In this paper we examine effects of the new risk and return paradigm in the Euro sovereign bond market on the tracking performance of 31 Euro zone sovereign debt exchange traded index funds (ETFs), during 2007-2010. Our results shed more light on the changing dynamics of Euro sovereign bond markets and the performance of ETFs. We contribute to the existing literature by examining the tracking performance using four different methods, including cointegration analysis. We further present novel evidence on the determinants of tracking errors using panel data analysis. To the best of our knowledge, this is the first study to examine

[3] The investigation is based on the large survey by the Edhec Institute (2009).

alternative tracking error measures and the determinants of the tracking performance of Euro zone sovereign bond ETFs.

Overall, ETFs underperform their respective benchmarks. All sample ETFs have statistically significant average (mean) standard (i.e. correlation based) tracking errors, at the 1% level of significance. The average underperformance during the sample period varies from 10 to 27 basis points annually and is thus of considerable economic interest. The underperformance peaked in 2008 (during Lehman crisis) and 2010 (during Greek sovereign crisis). Finally, ETFs which employ physical replication strategies exhibit higher degree of average underperformance.

Selection rules also contribute to different returns for two otherwise similarly structured indices. For example, sample funds with exposure to the riskiest sovereign issuers perform differently in comparison to otherwise similar funds that exclude the risky issuers. Our comparison of tracking errors based on the correlation and cointegration metric (i.e. residuals from the cointegrating regressions) further highlight the differences among ETFs. iShares funds, which replicate Barclays Term indices, exhibit the smallest, while Lyxor ETFs exhibit the largest correlation-based tracking errors. In contrast, Lyxor ETFs exhibit the best, and iShares funds the worst, tracking performance within a cointegration framework. The results also suggest that longer maturity segments have typically higher correlation based tracking errors but lower tracking errors obtained using cointegration metric.

The results of our panel data analysis document significant changes in the sample ETFs' tracking performance during the sovereign debt crisis. Our results also confirm that, as a result of the crisis, credit risk considerations are an important determinant of the ETFs' performance. We also find evidence for the importance of volatility of underlying indices and replication method for the tracking performance, irrespective of the tracking error metric. In an environment of widening sovereign credit default swap (CDS) spreads and divergent yield trends, understanding the credit quality of various issuers together with the selection rules of benchmark indices is, therefore, crucial for understanding ETFs' performance.

The remainder of the paper is organised as follows. In Section 2 we present characteristics of the European sovereign debt indices and ETFs. Section 3 presents relevant literature and motivates hypotheses. Section 4 describes the data and methodology. The results are discussed in section 5. Finally, we conclude in section 6.

2. Characteristics of Euro zone sovereign bond indices and ETFs

2.1. Euro zone sovereign bond indices

ETF indices are created by an institution which then licenses the index to an investment bank or a brokerage house. The dominant providers of Euro zone sovereign debt indices are: Barclays Capital, the International Index Company (IIC), EuroMTS, and Deutsche Borse. Between them they provide leading families of indices such as: Barclays Term, Markit iBoxx, eb.rexx and EuroMTS (See Table 1).

Insert Table 1 about here

Fixed income indices tend to be more complex than equity indices as bonds are typically not traded on organised exchanges. In constructing indices, index providers utilise different price sources including all relevant trading platforms currently operating in Europe. Typically, only standard coupon bonds that are redeemed on a fixed maturity date are eligible for inclusion into indices.[4] In order to be included in an index, the time to maturity of a bond has to be at least 1 year. Although the above index families target comparable maturity segments they have very different compositions and, thus, very different risk and return characteristics.

Barclays Capital, for example, pioneered the concept of a *term index*. In contrast other market indices, term indices have stricter inclusion criteria regarding both the original time to maturity and remaining time to maturity. They include only bonds with a remaining time to maturity near to their original time to maturity, rather than selecting all bonds in an index maturity range. As a result, term indices have similar yields, duration and risk/return characteristics to standard maturity-based indices and are more compact and liquid. The International Index

[4] Exceptions are Markit iBoxx benchmark indices which can also include standard discount (stripped) bonds.

Company (IIC) develops and runs the Markit iBoxx bond indices. A distinctive feature of iBoxx bond indices is multi-contributor real-time pricing (i.e. pricing that takes into account price information from multiple trading platforms) (See Table 1). iBoxx also calculates and publishes consolidated bond prices once per minute each trading day. iBoxx indices track the overall exposure to the Euro zone sovereign debt market. [5] The eb.rexx indices are calculated using the quotes from the Eurex Bonds platform.[6] Finally, EuroMTS indices track the overall exposure to the Euro zone sovereign debt market and are priced using real-time quotes from the MTS platform.

2.2. Euro zone sovereign bond ETFs

One of the key differences among ETFs is in the replication technique they employ.[7] It is either physical (in-kind) or synthetic (swap-based) replication. The fund manager of a physically-based ETF replicates its index through acquisitions of securities held in it. Consequently, the fund portfolio consists of all securities and incurs a relatively high transaction cost.

A synthetic ETF, on the other hand, lends its holdings (which may or may not constitute a sub-portfolio of a benchmark) to counterparty via a

[5] For more details about Markit iBoxx Liquid and Liquid Capped indices, see http://indices.markit.com/download/products/guides/Markit_iBoxx_EURLiquid_Guide.pdf

[6] For more details about eb.rexx indices, see http://www.dax- indices.com/EN/MediaLibrary/Document/ebrexx_L_3_8_e.pdf

[7] For more on differences between the replication methods, see BlackRock (2010).

collateralised repurchase agreement and then swaps the yield on that loan for the total return of the underlying index. The yield on the loan is based on LIBOR with or without a spread. [8]

Table 2 shows the summary aggregate country exposures of indices tracked by sample ETFs, by their providers.

Insert Table 2 about here

Barclays' indices predominantly focus on Germany, Italy, and France. Those 3 countries constitute typically three quarters of the overall index. The remainder is allocated to the Netherlands and to a smaller extent Spain. Spain is, therefore, the only so-called peripheral EU country with a weighting ranging from 7% to 13.3%. The iBoxx Euro Liquid Sovereign index caps any one country's weight to 20%. Consequently, the weightings outside Germany, France, and Italy are larger. For example, the allocation to Spain is around 20%, Belgium from 4.4% to 14.5%, Greece from 5.9% to 13.9%. Ireland is included in the 1.5-2.5 index (allocation of 3.8%). The eb.rexx index family includes only the most liquid standard coupon bonds issued by the German Government. They are, therefore, the least geographically diversified out of all the indices tracked by the sample ETFs. In terms of weighting, the db x-trackers (iBoxx Euro

[8] This is one of several possible swap-based index replication methods. Increased investor interest in the transparency and assessment of counterparty risk has made the evolution of these structures. More recently, swap-based funds can have collateralised or uncollateralised counterparty exposure and a single counterparty or multiple counterparties. As in December 2010, ETFs from the iShares family are based on full physical replication while db x-trackers and Lyxor funds are based on synthetic replication, as described in the text.

Sovereign) index and Lyxor (EuroMTS) indices do not differ substantially. They both include riskier countries such as Ireland, Greece, Portugal, and Belgium in similar proportions.

3. Literature and hypotheses

Tracking errors have 3 main sources: transaction costs related to constructing the indexed portfolio (including management fees), differences in composition of the index fund and the index itself, and trading noise. Trading noise is associated with timing differences when the underlying index is published continuously while ETF trades less frequently (Flood, 2010b). Another example of the noise are discrepancies between the prices used by the ETFs constructing the index fund and actual transaction prices (Fabozzi, 2000). A smaller tracking error means better tracking performance of the fund with respect to its underlying index. The main task of an ETF manager is, therefore, to find an optimal tradeoff between the closeness of index replication and the cost of replication.

Previous studies of the Europe-based ETF tracking errors primarily focused on equity ETFs. Blitz et al. (2010), for example, compare Europe-listed index mutual funds and index ETFs that offer exposure to global equity markets. The authors report that, on average, ETFs underperform their benchmarks.[9] Rompotis (2008) and Milionas and Rompotis (2006)

---

[9] For research on various aspects and characteristics of equity ETFs outside Europe see: Potreba and Shoven (2002) – US; Gallagher and Segara (2005) - Australia, and Jares and Lavin (2004) –

examine performance of German and Swiss equity index ETFs. Both studies report a slight underperformance of ETFs relative to underlying indices. They also report that tracking error is positively related to risk (measure by standard deviation of returns on target indices) and management fees. In this study, we use a weekly percentage change in the level of respective indices as proxy for the underlying risk. We conjecture that this is more appropriate measure since it captures both frequency and magnitude of short term changes in the underlying indices. Thus, indices that change more frequently, and by larger percentages, would be more difficult to track. Based on the previous evidence for equity ETFs, we test the following hypothesis:

H1: Euro zone sovereign ETFs underperform their underlying bond indices

H2: Tracking performance is negatively associated with volatility of tracked indices

Alexander (1999) highlight the importance of examining tracking performance based on both correlation and cointegration. Although related, the concepts of correlation and cointegration are quite different. For example, static correlations measure short-term co-movements and, thus, exhibit significant instability over time. In contrast, cointegration measures long-run co-movements. Two time series can be cointegrated even if static correlations between them are low (Alexander, 1999). Consequently, investment strategies based only on volatility and the

---

Japan and Hong Kong. For comparison of ETFs and index mutual funds that track the same stock indices (e.g. S&P 500) see: Delva (2001) and Kostovetsky (2003).

correlation of returns suffer from a number of drawbacks, especially when applied to a passive investment framework. For example, if the tracking error follows a random walk process, the portfolio needs to be frequently rebalanced in order to prevent diverging significantly from its benchmark (Alexander, 1999). The minimization of the tracking error with respect to an index, however, contains a significant amount of noise and may result in a portfolio that is unstable in volatile market circumstances (Alexander and Dimitriu, 2004). We, therefore, test the following hypothesis:

H3: Correlation and cointegration based TE metric may provide different ranking (assessment) of the tracking performance

The recent sovereign debt crisis has changed the dynamics of Euro bond markets. For example, the average euro sovereign spreads widened sixfold in the 2008-10 period, while spreads' volatility increased twentyfold (Nomura, 2011). [10] The increase in credit risk and volatility of credit defaults swap premiums makes sovereign bonds similar to their corporate counterparts. Whilst before the crisis risk of Euro sovereign bonds was almost entirely attributable to yield curve risk, during 2008-2010 peiod the increase in risk was predominantly due to the spread volatility (Nomura, 2011). At the same time, the average pair-wise correlations between peripeheral sovereign spread changes reached

---

[10] Interestingly, during the same period time both corporate credit spreads and volatility increased approximately threefold (Nomura, 2011).

historical high of approximatelly 80%, thus reducing potential benefits of diversification.[11]

Furthermore, price elasticity of deficit differentials (in terms of GDP) has increased 3-4 times while price elasticity of differentials in the level of debt (as a fraction of the GDP) has increased around 7-8 times during the post-Lehman crisis period (Schuknecht et al., 2010). In the, post-Greek debt crisis period, markets continue to penalise fiscal imbalances much more strongly (Attinasi et al., 2009). The above developments further increased investors' interest in sovereign debt ETFs. They, however, also radically changed the nature of risk in the Euro sovereign debt market. We conjecture that these developments changed ETFs' tracking performance with respect to the pre-crisis situation. Thus,

H4: Association between ETFs and underlying bond indices (i.e. tracking performance) changed during the period 2008-2010

Given the new risk return paradigm in the Euro sovereign bond market, more sophisticated approach is required for assessment of the determinants of the performance of the sovereign ETFs. For example, the differences in the tracking performance are expected to be associated with the composition of indices, especially to their exposure to the (higher risk) peripheral EU countries. Norden and Weber (2009) examine intertemporal relationships between CDS, bond, and equity markets. They report that CDS markets reflect default-risk related information earlier than bond

[11] The average pair-wise correlations estimated across a rolling 52-week window (Nomura, 2011; p.82).

market and that bond spreads adjust to CDS spreads. We conjecture that this may create another source of tracking errors. The testable hypothesis is, therefore, a significant association between ETF's tracking errors and volatility of sovereign credit default swap (CDS) spreads. Since the spread changes tend to be log-normally distributed, the dynamics of spread volatility is driven by those of spread levels and of the underlying volatility of proportional spread changes.[12] Thus,

H5: CDS spread volatility (consisting of both levels and underlying volatility) has become an important determinant of tracking performance

Finally, we control for different fund characteristics such as replication method, underlying maturity, and bid-ask spread for ETFs. The bid-ask spread, for example, has been identified as an important determinant of the tracking performance of equity ETFs (Rampotis, 2008). In the context of our study, we expect that bid-ask spread proxy for liquidity of premium associated with different ETFs.[13] We further expect that longer maturity segments exhibit higher yield curve risk. Given a very high correlation between size and maturity of our sample ETFs, the maturity is also controlling for the size of ETFs. We further expect that so called synthetic ETFs provide a better long term replication than in-kind ETFs

[12] For more on importance of examining both levels and changes in spreads, see Nomura (2011).
[13] It is important to note that ETFs' liquidity should not be proxied by their trading volume since they are not necessarily related. There is also some anecdotal evidence that average expense ratio tend to be higher for ETFs with lower liquidity. See, www.etftrends.com.

due to higher transaction costs associated with physical replication method.[14]

4. Data and methodology

4.1 Data and sample characteristics

We examine 31 ETFs of leading European providers: iShares (track Barclays Term, Markit iBoxx Liquid Capped and eb.rexx German Government indices), Lyxor Asset Management (track EuroMTS indices) and Deutsche Bank db x-trackers (track Markit iBoxx benchmark indices).[15] The above mentioned ETF providers were ranked as top one, two, and three European ETF providers, respectively, in 2008.[16] All sample indices belong to the class of total return indices (i.e. indices where all coupon payments are re-invested). Our sample ETFs from the iShares family (Barclays, iBoxx Liquid Sovereign, eb.rexx) are based on the physical replication while db x-trackers (iBoxx Sovereign) and Lyxor funds (EuroMTS) are based on the synthetic replication method.

The principal source of our data is Bloomberg and the official websites of ETF and index providers. We use daily closing ETFs; Net Asset Values (NAV), bid-ask spreads, and daily weights for different index constituents, for the period between January 2007 and December 2010. For

[14] This is consistent with some anecdotal evidence provided by professional investors (see Michalik, 2011). It is, however, worth noting that swap based replication method creates a counterparty risk.
[15] EuroMTS indices tracked by our sample Lyxor ETFs are also known as EMTX indices.
[16] Worldwide, iShares were ranked as number one, Lyxor number four, and db.trackers number five (Bloomberg, 2008).

funds which did not exist on January 2007, the date of their inception is used as the first date of the corresponding time series.[17] Given that sample ETFs are listed on multiple exchanges, for consistency reasons, we use data from the German listings (Frankfurt exchange).

Assets under management (AUM) of our sample funds are of the order of hundreds of millions of Euros. ETFs tracking longer maturity segments (over 10 years of maturity) tend to be larger than their counterparts tracking shorter maturity indices. For instance, in the May of 2010, the well established short-maturity funds such as iShares eb.rexx 1.5-2.5 and Lyxor EuroMTS 1-3 had an AUM of around € 1.3 and € 1.03 billion, respectively. On the other hand, longer maturity funds such as db iBoxx 10-15, 15+, 25+, and iShares Markit iBoxx € Liquid Sovereigns Capped 10.5+ had an average AUM of roughly € 30 and €22 million, respectively.

4.2 Methodology

4.2.1 Tracking error – approach based on correlation of returns

Tracking error is the single most important factor in the managing and analysis of an index fund performance (Pope and Yadav, 1994). The ETFs performance may differ from the performance of the underlying index due to imperfect replication strategy and market frictions that affect EFTs and not necessarily targeted indices. In order to determine how closely an ETF tracks the targeted index, we calculate tracking error (TE1) as the active return,

[17] All db x-tackers ETFs started trading in May and June 2007, except for the db short iBoxx index, which started trading in May 2008. Funds iShares Barclays 5-7 and 10-15 started trading in April 2009, while Lyxor EuroMTS 15+ started trading in June, 2007.

i.e. as the difference in returns between the ETF and the index at the end of a certain period of time. TE1, therefore, measures the difference between the return of a fund and its underlying index.[18]

$$TE1 = \left(r_p - r_b\right) \qquad \textit{Equation (1)}$$

Where r_p is the return on the ETF and r_b is the return on the underlying index.

The goal of a passive investment strategy is to gain exposure, as accurately as possible, to all index characteristics and not just to match the value at the end of the investment horizon. One way to do this is to compare the volatilities of the fund with that of the benchmark. However, the comparison of volatilities alone would ignore co-movements between indices and ETFs. Furthermore, the variability of total returns may not be symmetric in rising and falling markets (Fabozzi, 2000). Having this in mind, tracking error (TE2) is commonly defined as standard deviation of the difference between the return on the portfolio and that of the benchmark: [19]

$$TE2 = stdev\left(r_P - r_B\right) \qquad \textit{Equation (2)}$$

[18] The difference is often referred to as active returns.
[19] See Alexander (2008), Bacon (2008), and Martellini et al. (2003) for more details on methodological issues regarding measurement of tracking performance.

Equation (2) is the most frequently used performance measure of index funds among practitioners. Although it describes variability in active returns, TE2 provides no information on a fund's under- or over-performance vis-à-vis the benchmark index. It ranks equally both the positive and negative active returns of the same magnitude. Thus, as a performance measure TE2 is more appropriate for tracking (index) funds and less appropriate for active funds.

Equation (2) also measures co-movement of portfolio returns with that of a benchmark:

$$TE2 = \sqrt{\sigma_P^2 + \sigma_B^2 - 2\sigma_P \sigma_B \rho_{P,B}} \qquad \textit{Equation (3)}$$

In this expression, σ_P is the standard deviation of ETF returns, σ_B is the standard deviation of underlying index returns, and $\rho_{P,B}$ is the correlation between ETF and index returns. Clearly, the higher the correlation the lower the tracking error measure TE2, *ceteris paribus*.

It is worth mentioning that TE2 tends to be good in detecting the trading noise while it largely ignores the bias introduced by fund management fees. While trading noise often causes an ETF to close at a slight premium, management fees tend to produce consistent daily underperformance compared to the benchmark index. The accumulated net effect of these daily deviations is much more important for long term investors than for short term investors.

4.2.2 Tracking error - OLS approach

The following OLS regression of ETFs' returns on index returns is also used as a measure of tracking performance in previous literature (Pope and Yadav, 1994):

$$r_p = \alpha_i + \beta_i r_b + \varepsilon_p \qquad \textit{Equation (4)}$$

Where r_p is the return on the ETF, r_b is the return on the relevant index, and ε_p is the error term. The standard error of the above regression is another measure of tracking error (TE3). In the case of ETFs pursuing a passive investment strategy, alphas are expected not to be statistically different from zero. In a passive investment framework, betas are expected not to be statistically different from one and R^2s are expected to be very high.[20] Pope and Yadav (1994), however, highlight important problems associated with TE3. First, if beta is not exactly equal to 1, TE3 may give different ranking to TE2. Second, the OLS approach may overestimate the tracking error if the relationship between ETF and index returns is not linear. Rankings provided based on TE2 and TE3 are expected to be very similar.

4.2.3 Tracking error – cointegration framework

[20] Cresson et al. (2002) suggest R^2 from equation 4 as another, more straightforward, measure of tracking performance.

As all sample funds passively track respective indices, we expect the series of sample ETFs to be cointegrated with the respective benchmark indices. In the context of cointegration, the tracking error is the residual from the cointegrating regression (i.e. cointegrating vector).[21] We define this tracking error as TE4. The more stationary the tracking error (TE4), the greater the cointegration between the ETF and the respective index (Alexander, 1999). In order to examine the tracking errors within cointegration framework, we fit the following Vector Error-Correction (VEC) model for our sample ETFs and their underlying indices: [22]

$$\Delta NAVp_t = \alpha_1 + \sum_{i=1}^{m_1} \beta_{1i} \Delta NAVp_{t-i} + \sum_{i=1}^{m_2} \beta_{2i} \Delta NAVb_{t-i} + \gamma_1 z_{t-1} + \varepsilon_{1t} \qquad \textit{Equation (5)}$$

$$\Delta NAVb_t = \alpha_2 + \sum_{i=1}^{m_3} \beta_{3i} \Delta NAVp_{t-i} + \sum_{i=1}^{m_4} \beta_{4i} \Delta NAVb_{t-i} + \gamma_2 z_{t-1} + \varepsilon_{2t} \qquad \textit{Equation (6)}$$

Here, NAV_p and NAV_b are daily log NAV values for ETFs and underlying indices, z is the cointegrating vector, Δ is the first difference operator, and the lags, lengths and coefficients are determined by OLS regression.

The coefficients β_{2i} and β_{4i} are adjustment parametars. They provide inferences regarding causality flows and dynamic correlations between ETFs and index returns. For example, if the coefficient β_{2i} in equation 5 is found to be statistically significant, then the turning points in $NAVb$ will

[21] In the context of this study we use the cointegration framework to examine the tracking error of sample ETFs. For other applications of cointegration analysis in portfolio management, see Alexander (1999).
[22] The model is adapted from Alexander (1999).

lead turning points in $NAVp$. The opposite applies for the coefficient β_{4i} in equation 6.

There are several different frameworks for the estimation of the cointegrating systems. We follow the methods developed by Johansen (1988; 1995). First, to test for cointegration and fit the cointegrating VEC model we need to specify a number of lags. We determine the number of lags based on a Hannan-Quin information criterion (HQIC), Schwartz Bayesian information criterion (SBIC), and sequential Likelihood Ratio (LR) methods. We then use the Johansen test (JT) to determine the number of cointegrating equations between daily log NAV values for ETFs and relevant indices together with the results of the VEC model.[23] After determining that there is a cointegrating equation between log levels of ETFs and respective indices, we estimate the parameters of a bivariate cointegraiting model based on maximum likelihood (ML) estimations.[24] The sample ETFs that are cointegrated with the underlying indices are, therefore, those with better tracking performance. The information about the models' fit and residauls can be used to compare tracking performance among sample ETFs. For example, the more stationary the tracking error (TE4) the better the tracking performance of the sample ETFs.

[23] In a bivariate case like ours, at most one cointegrating vector is expected, otherwise the original series would have to be stationary (Alexander, 1999).
[24] Based on graphical examinations of the original series, we selected estimations that allow for a linear trend.

4.2.4 Determinants of tracking errors

We start by examining the possible determinants of tracking errors (TE2 and TE4) such as: replication method (i.e. synthetic vs. physical-backed replication), ETFs' maturity, index volatility, and ETFs' bid-ask spreads, within a panel data framework:

$$TE_{it} = \alpha + \beta_1 BID - ASK_{it} + \beta_2 LNTERM_{it} + \beta_3 SYNTHETIC_{it} + \beta_4 WINDEX_{it} + \varepsilon_{it}$$

Equation (7)

Here, TE_{it} is $TE2_{it}$[25] or $TE4_{it}$; BID-ASK is the daily bid-ask spread for ETFs; LNTERM is the natural logarithm of the maturity of the sovereign bonds included in the underlying indices (in months);[26] WINDEX is the weekly percentage change of respective indices; SYNTHETIC is a categorical variable that equals to 1 if the ETF adopts a synthetic rather than a physical replication and zero otherwise.

We then consider how recent changes in the European sovereign bond markets affected the tracking performance of the sample ETFs. First, we conduct Quandt-Andrews Breakpoint tests for equation (4), in order to identify any structural breakpoints in the relationship between ETFs' and index returns.

---

[25] TE2 is based on daily returns.

[26] Given the range of maturities for the sovereign bonds included in the underlying indices, each of the sample ETFs and their corresponding benchmarks were in a first step ranked. For example, the ETFs with the shortest maturity bracket (1-3 years) received ranking equal to 1. The ETFs with the next maturity bracket (3-5 years) received ranking equal to 2, etc. The ranking was than multiplied by 12 months. The resulting number of months was used as a proxy for the average maturity of the sovereign bonds included in the underlying indices.

For 18 sample ETFs structural breaks were identified during the period from June 2008 to October 2009, which is the period of Lehman crisis.[27] In one case, the structural break was identified during the peak Greek sovereign crisis, in April 2010.[28] For 3 sample ETFs, the structural break was identified during the period from January to August 2009. Finally, in two cases no structural breaks were identified at reasonable levels of statistical significance.

We, therefore, control for these two periods by introducing categorical variables LCRISIS (during the period from 1st September 2008 to 30th September 2009) and GCRISIS (from 19th April 2010 to 23rd December 2010).[29]

Second, we construct new variables to capture both spread leavels and the underlying volatility of proportional spread changes. The variables are constructed based on daily values for 5-year Credit Default Swap (CDS) spreads for euro sovereigns included in different bond indices. We calculate weighted average of CDS spreads for each of the bond indices tracked by our sample ETFs, using daily weights and CDS spreads for the index constituents. The weighted averages represent spread levels. The underlying volatility is then estimated as a weekly percentage change of the weighted average CDS spread. Finally, we control for the eb.rexx

[27] In all of the cases the results of Quandt Andrew tests were highly statistically significant.
[28] These findings are consistent with our results presented in Figures 1 and 2.
[29] It is worth noting that apart from structural breaks during the Lehman crisis have also numerous structural breaks during the sample period, highlighting their time-varying properties. The simple categorical variables introduced above, thus may not be most appropriate in describing the switching process for Lyxor funds.

family since ETFs from this family consist of only German government bonds:[30]

$$TE_{it} = \alpha + \beta_1 BID - ASK_{it} + \beta_2 LNTERM_{it} + \beta_3 SYNTHETIC_{it} + \beta_4 WINDEX_{it} + \ldots$$
$$\ldots + \beta_5 LNWCDS_{it} + \beta_6 WPCDS_{it} + \beta_7 LCRISIS_{it} + \beta_8 GCRISIS_{it} + \beta_9 EBREXX_{it} + \varepsilon_{it}$$

Equation (8)

Here, LNWCDS is the natural logarithm of weighted average CDS spreads; and WPCDS is the weekly percentage change in the level of weighted average CDS spreads.[31]

5. Results

5.1 Volatility and correlation of sample ETFs and respective bond indices

Table 3 presents the results for the volatility of the sample ETFs and respective indices together with their correlations during the sample period. All ETFs exhibit a similar volatility to the volatility of underlying indices. The correlation between ETFs and indices is very high (ranging from 0.97 to 0.99) except for Lyxor's ETFs with the correlation of 0.51. The average correlation between sample ETFs and respective bond indices has increased for Lyxor and iBoxx Liquid families, but dropped for ETFs from eb.rexx family, in 2008. The average correlation for other families of sample ETFs remained the same. In the aftermath of Lehman crisis the

[30] The data on weights for individual constituents was not available for the db x-trackers family of ETFs and they were not examined in equation 8.
[31] All other variables are defined as in equation 7.

average correlation dropped, resulting in lower average correlations for iBoxx Liquid, eb.rexx and Lyxor families, in 2009. Interestingly, the average correlation for Barclays family increased while the correlation of iBoxx Sovereign remained the same as in 2008. This is consistent with the results in Drenovak and Urošević (2011) obtained for time series up to May 2010. In 2010, the average correlation dropped for Barclays and Lyxor families but increased for all other sample families returning close to 2007 levels.

Insert Table 3 about here

Overall, db-trackers and Lyxor ETFs, together with their respective indices, exhibit higher average volatility during the sample period compared to other sample ETFs. In Figure 1 we also present the evolution of volatilities for the sample ETFs targeting 10+ year maturities. The volatility increased significantly from the latter part of 2008, which corresponds with the beginning of the financial crisis. Since reaching its peak in the spring of 2009, volatility has exhibited a decreasing trend until April 2010. The above results confirm changing dynamics in the market for Euro sovereign bonds and are broadly consistent with the earlier identified structural breaks during the sample period.

Insert Figure 1 about here

5.2 ETFs' performance and tracking error (TE1)

The ETFs annual returns and benchmark adjusted performance of ETFs is presented in Table 4. Overall, sample ETFs' returns increased significantly in 2008. The high returns reflected an increase in popularity of sovereign debt as investors sought alternatives to equity markets and as interest rates continued to decline. The situation, however, changed with signs of a financial crisis in late 2008. Consequently, the returns for all ETFs dropped sharply in 2009, reflecting an increase in CDS spreads and speculation that one or more countries could be forced to leave the single currency. More recently, we have witnessed a return to higher returns similar to those during the pre-crisis period.

Insert Table 4 about here

In terms of the specific funds' performance, Barclays Term and db-x trackers funds performed better than their counterparts. iBoxx Liquidity ETFs were the worst performers among the sample funds. Notably, they have the highest weighting for Spain, Greece, and Belgium compared to all other sample ETFs.[32]

Overall, ETFs underperformed their respective indices. The average underperformance during the sample period varies from 10 basis points

[32] It is worth mentioning that CDS spreads do not affect bond yields of different maturities to the same extent. First, referred CDS spreads are typically those for a 5-year maturity, as the most liquid maturity while bonds may have shorter/longer maturities. Second, there is normally a basis risk equal to the differences between CDS spreads and bond yield spreads (Financial Times, 2009).

(Lyxor) to 27 basis points (iShares - Barclays).[33] The underperformance is more pronounced for funds which employ physical replication (i.e. iShares). The level of the ETFs underperformance increased sharply in 2008 and then dropped in 2009, only to increase again during 2010.

5.3 Correlation based tracking errors (TE2)

The results of the sample ETFs' tracking errors (TE2) are presented in Table 5. All sample ETFs have statistically significant average (mean) tracking errors at the 1% level of significance.[34] This result is robust to the use of monthly instead of daily NAV series. Overall, iShares funds which replicate Barclays Term indices exhibit the smallest while Lyxor ETFs exhibit the largest values of TE2. This finding is consistent with earlier reported differences in correlations between ETFs and respective indices. The highest TEs, thus, are associated with the lowest correlation with the underlying index.

Insert Table 5 about here

The results also suggest that ETFs tracking higher maturity indices have typically higher levels of tracking error. This is the case for all sample ETFs except for db iBoxx Sov 5-7. There are also some differences in the way TE2s changed during the sample period. iBoxx Liquidity and eb.rexx, for example, exhibited the highest TE2 in 2009. This is consistent with high weightings for Greece (in iBoxx liquidity) and Germany (in eb.rexx), the

[33] Lyxor ETFs actually overperformed respective benchmarks in 2007.

[34] Unreported results for one sample Wilcoxon test for median are economically and statistically consistent with the reported results for T tests.

two countries with extremely volatile interest rates during 2009. Barclays, db, and Lyxor indices, however, exhibited the highest TE2 during 2008.

Table 6 presents the results for tests for the differences in mean and median TE2s for different families of ETFs. The results confirm the superior tracking performance of iShares Barclay's family followed by db x-trackers, as measured by TE2. The difference in mean and median TE2 between Barclay's family and the rest of the sample is statistically significant at the 5% level of significance, or better. The difference between mean TE2 for db x-trackers and eb.rexx families is not statistically significant, confirming the similar performance of eb.rexx and db x-tracker funds.

Insert Table 6 about here

As can be concluded from Tables 3-6, iShares funds which replicate Barclays indices have the lowest average volatilities, lowest average TE2 values and the highest correlation coefficients with benchmark indices. This could be attributed to a compact index structuring approach applied for Barclays indices. In addition, iShares employ full physical replication strategy. The physical replication involves taking possession in most or all of the positions of the benchmark portfolio. In this case, fund and benchmark returns are highly correlated (they would be identical if expenses and income from other activities are adjusted for). This leads to a low variability in active returns and, therefore, to lower TE2.

In comparison with iShares funds, db x-tracker and Lyxor funds, track indices have more constituents and more complex country exposures. In addition, they, employ swap-based instead of physical replication. Our analysis shows that Lyxor funds exhibit a considerably higher average TE2 compared to those of the other two providers. Such a high TE2 could be explained by a lower correlation between the fund and benchmark returns (see Table 3) rather than by funds' characteristics that are similar to db x-trackers.[35] The different performance of Lyxor funds is also evident from the results presented in Figure 2, which depicts patterns of three month average TE2s for the sample ETFs. The figure also highlights the positive association of TE2s with levels of volatility.

Insert Figure 2 about here

5.4 Results of the OLS model (TE3)

The results presented in Table 7, suggest that only 3 out of 31 sample ETFs have alphas statistically different from zero. The estimations for all beta coefficients are statistically different from 1 for all but 3 sample funds. Lyxor family is a clear outlier with an average beta of 0.53 confirming that ETFs from this family depart from full replicating strategy. The coefficient

[35] Obviously, with a swap replication strategy it is the swap contract that defines characteristics of fund replication. So, to compare the performance qualities of such funds it would be crucial to know the details of the swap contract and provisions that determine the reference basket that is the basket of securities with the fund. In cases when the fund manager takes part in swap collateral lending one should also analyse the structure of the collateral basket. However, we currently do not have access to such data.

beta for all is below 1 for all but two sample funds, indicating that the sample's ETFs are more conservative than their respective benchmarks.

Insert Table 7 about here

Regarding R^2, the Lyxor family is again a clear outlier with an average of R^2 only 30%. For all other sample ETFs, R^2 ranges between 93% and 99%, indicating a very good regression fit. The iShares family exhibits the highest average R^2 of 98%. Reported values for TE3, measured by standard deviation of residuals, indicate the same ranking of sample funds as is it was case when we employed TE2 measure.

5.5 Results of cointegration analysis (TE4)

Table 8 presents the results of Johansen's test for cointegration and our VEC model. We present the number of lags and Johansen's test (JT) statistics for each of our sample ETFs. The normalised cointegrating vector, log likelihoods, are presented for the sample ETFs with a cointegrating vector. Dickey-Fuller test statistics for residuals of the cointegration regression was presented and used for ranking of sample funds' tracking performance (TE4).

Overall, our results suggest strong support for cointegration equations 5 and 6. Based on the results of the Johansen test for cointegration we strongly reject the null hypothesis of no cointegration in all but five ETFs. Four out of five funds that show no evidence to support the existence of a cointegration equation are from the iShares family (1-3, 3-5, 7-10, and 15-30). In contrast, the log likelihoods, Johansen test statistics, and Dickey-Fuller test statistics (DF) for residuals, indicate a very good model fit for the iBoxx and the Lyxor ETFs, followed by the eb.rexx and the db x-trackers families. [36] Within each family of sample ETFs the tracking error proxied by DF test statistics tend to be lower for indices with longer maturities. The coefficients for indices in the cointegration equations are negative and highly statistically significant for all ETFs, suggesting rapid adjustment toward equilibrium.

Adjustment parameters in our bivariate model tell us more about dynamic correlation and causalities between returns of ETFs and underlying indices. For Lyxor family the coefficent *beta2* is negative and statistically significant while *beta4* is not.[37] This suggests the existence of causal flows from indices to Lyxor ETFs. For example, when the average *NAVp* are too high, it quickly falls back toward the *NAVb* level. The results are similar for sample ETFs from the db x-trackers family.[38] The adjustment pareametars for sample ETFs from the eb.rexx family, however, suggest causal flows from ETFs to underlying indices. The results for the iBoxx family is less conclusive suggesting causal flows in both directions, from indices to ETFs, and from ETFs to indices. Unreported results for iShares family suggest lack of statistical significance for the adjustment

---

[36] It is worth noting that Lyxor funds exhibit the highest while iShares funds exhibited the lowest TE2 in Table 5.

[37] For reasons of brevity we did not present complete results with adjustment parameters for all models. The results are available upon request.

[38] It is important to note that these two fund families follow synthetic replication strategy.

parametars, confirming previously reported reelatively poor fit of the cointegrating model for these funds.

The results of our cointegration analysis, combined with the results presented in Table 5, suggest that ETFs from the iBoxx, the Lyxor, and the eb.rexx families may have adopted a long term view attempting to reduce tracking errors by taking into account common long term trends, whilst other providers tend to focus more on tracking errors based on short term correlations.

Insert Table 8 about here

5.6 Determinants of tracking errors

The correlation matrix of all variables (from equation 8) is presented in Table 9. The model is estimated as random effects GLS regression, with a robust estimator of variance adjusted for clustering on fund.[39] The explanatory variables do not exhibit high correlation suggesting that multicolinearity is not a problem in our panel regression model. The highest (positive) correlation (0.47) is between LNWCDS and a categorical variable GCRISIS. This is consistent with a significant increase in CDS spreads, especially for peripheral euro sovereigns during the Greek sovereign crisis in 2010.

[39] Unreported results for Breusch and Pagan Langrangian Multiplier test for random effects, and Hausman test for systematic differences in coefficients provide justification for the selection of the random (over fixed) effects model.

Insert Table 9 about here

The results of the panel models are presented in Table 10. The results for equation 7 indicate the positive and statistically significant association of tracking error (TE2) with BID-ASK, SYNTHETIC, and WINDEX (Panel A). The results with TE4 as a dependent variable, in Panel A, are consistent for SYNTHETIC, and WINDEX. Sample ETFs with the synthetic replication method tend to have higher TE2 and TE4 compared to their counterparts. As expected, the higher variability of underlying indices result in higher TE2 and TE4. The coefficient for LNTERM is still with positive sign, but this time is statistically significant at the 1% level. The difference between the regressions for TE2 and TE4 is in the sign of BID-ASK variable. The coefficient for BID-ASK is positive and statistically significant at the 5% level in the model for TE2, but negative and statistically significant at the 1% level in the model for TE4. Overall, our panel data model exhibits higher explanatory power for the TE2 model for TE2 (overall R^2 of 17.6%) compared to the model for TE4 (overall R^2 of 1.3%).

Insert Table 10 about here

With the inclusion of variables associated with levels and changes in CDS premiums and categorical variables for crisis periods, the explanatory power of both models increases significantly (Panel B).[40] For example, R^2

[40] The unreported results of the estimations using the maximum likelihood method are economically and statistically consistent with the results reported in Table 10.

is 57.50% and 2.97% for models TE2 and TE4, respectively. The coefficients for GCRISIS and LCRISIS are positive and statistically significant at a 1% level both for TE2 and TE4, confirming our conjecture of poorer tracking performance during the period of the credit crisis. The coefficient for WINDEX is positive and statistically significant, both for TE2 and TE4, indicating a positive association of tracking errors and volatility of respective indices. The coefficient for LNWCDS is negative and highly statistically significant at the 1%, both for TE2 and TE4 models. The difference between the results for TE2 and TE4 is again related to BID-ASK. The coefficient for BID-ASK is positive and insignificant for TE2, but negative and statistically significant at the 5% level in the model for TE4. The coefficient for LNTERM has the same (positive) sign but it is higly statistically significant only in the model for TE4.

We tried an alternative specification of the models, without LNWCDS. With the new specification for TE2 model, the coefficient for GCRISIS becomes negative and statistially significant at the 1% level, while the coefficients for all other variables remain similar to those presented earlier thus suggesting improvements in tracking performance in 2010. The documented recent improvements in tracking performance are consistent with the results presented in Table 5. For TE4 model, LCRISIS is not statistically significant any more. The only other difference is change of the coefficent for EB.REXX from negative to positive sign (statistically significant at the 5% level).

6. Summary and conclusions

We examine the tracking performance of 31 Euro zone sovereign debt exchange traded index funds (ETFs), during 2007-2010. The tracking performance was examined using traditional, OLS, and cointegration tracking error models. We also provide evidence for changing dynamics and new risk and return paradigm in the Euro sovereign bond market.

Overall, ETFs underperform their respective benchmarks, thus lending support to our hypothesis 1. There are, however, some important differences across families of ETFs. iShares funds, which replicate Barclays Term indices, exhibit the smallest while Lyxor ETFs exhibit the largest correlation-based tracking errors. In contrast, within a cointegration framework, Lyxor ETFs (together with iBoxx) exhibit the best while iShares ETFs exhibit the worst tracking performance. Our results confirm consistency of TE2 and TE3 tracking measures for the sample funds. However, the tracking error measures based on correlation between returns provide different ranking to one provided by the cointegration based metric thus providing support for our hypothesis 2.

The results also suggest that higher maturity segments have typically higher levels of tracking errors measured by both correlation and cointegration metric. We also show that index selection rules can result in the significantly different performance of two, at first glance, similarly structured indices that use same replication methods. For example, funds with exposure to the riskiest sovereign issuers (e.g. iBoxx Euro Liquid) exhibit different performance in comparison with funds that exclude the risky issuers (Barclays Term and eb.rexx). [41]

Our panel data analysis reveals statistically significant changes in the sample ETFs' tracking performance during 'Lehman' and 'Greek' crisis. Our results also confirm that, as a result of financial crisis, credit risk considerations are increasingly important for performance of sovereign bond ETFs. The above results lend support to our hypotheses 3 and 4. We also find evidence for the importance of volatility of underlying indices and replication method for the performance. In an environment of widening sovereign credit default swap (CDS) spreads and divergent yield trends, understanding the credit quality of various issuers together with the selection rules of benchmark indices is, therefore, crucial for understanding ETFs' performance.

[41] This is consistent with recent developments in the EFT market. For example, provider of Markit iBoxx indices excluded Greek government sovereign bonds from its indices after their rating fallen below investment grade, in early 2011.

REFERENCES

Alexander, C., 2008. Market Risk Analysis: Practical Financial Econometrics, John Wiley & Sons, England.

Alexander, C., 1999. Optimal Hedging Using Cointegration, from Mathematics of Finance, published in Philosophical Transactions: Mathematical, Physical and Engineering Sciences, Vol. 357, No.1758, 2039-2058.

Alexander, C. and A. Dimitriu, 2004. A Comparison of Cointegration and Tracking Error Models for Mutual Funds and Hedge Funds, ISMA Centre Discussion Papers in Finance, March.

Amenc, N., Goltz F., 2009. The Way Ahead for Exchange-Traded Funds: Results from a European Survey, Journal of Alternative Investments, Vol. 12 Issue 1, 50-54.

Attinasi, M., Checherita, C., Nickel C., 2009. What Explains the Surge in Euro Area Sovereign Spreads During the Financial Crisis of 2007-09? , ECB Working paper No. 1131, December.

Bacon, C., 2008. Practical Portfolio Performance Measurement and Attribution, John Wiley & Sons, England.

Barclays Global Investors, 2009, ETF Landscape Industry Review, August 2009.

BlackRock, The Evolution of Swap-based ETFs, iShares, September 2010.

Blitz, D., Huij, J., Swinkels L., 2010. The Performance of European Index Funds and Exchange-Traded Funds, European Financial Management, forthcoming.

Bloomberg, 2008, Overview of the ETF Industry and Implementing Asset Allocation, Report.

Cresson, J., R. Cudd, and T. Lipscomb, 2002, 'The Early Attraction of S&P Index Funds: Is Perfect Tracking Performance an Illusion'? Managerial Finance, 28, 1-8.

Delva, I.W., 2001.Exchange Traded Funds Not for Everyone, Journal of Financial Planning, 14, 110-124.

Drenovak, M., B. Urošević, 2011. Exchange-traded funds of the Eurozone Sovereign debt, Economic Annals, Vol.60, 31-60, forthcoming.

Fabozzi, F.J., Bond market analysis and strategies, 4th ed., Prentice Hall, 2000.

Financial Times, Sovereign Default Risk and European Bond ETFs, 10th March 2009.

Flood, C., iShares considers expanding fixed income ETF range in Europe, Financial Times, 9th August 2010a.

Flood, C., ETF tracking errors can 'mislead', Financial Times, 5th October 2010b.

Gallagher, R.D. and R. Segara, 2005. The Performance and Trading Characteristics of Exchange Traded Funds, Working paper – The University of New South Wales.

Jares, E.T. and A.M. Lavin, 2004; Japan and Hong Kong Exchange Traded Funds (ETFs): Discounts, Returns, and Trading Strategies, Journal of Financial Services Research, 25, 57-69.

Johansen, S., 1988. Statistical Analysis of Cointegration Vectors, Journal of Economic Dynamics and Control, Vol. 12, 231-254.

Johansen, S., 1995, Likelihood-Based Inference in Cointegrated Vector Autoregressive Models, Oxford: Oxford University Press.

Longstaff, F., 2004, The Flight-to-Liquidity Premium in U.S. Treasury Bond Prices, Journal of Business, Vol. 77, No.3, 511-526.

Kostovetski, L., 2003. Index Mutual Funds and Exchange Traded Funds, Journal of Portfolio Management, 29, 80-92.

Michalik, T., Synthetic Method Wins on Tracking Precision and Cost, FT special report on ETF 7th February 2011.

Martellini, L., P. Priaulet and S. Priaulet, 2003. Fixed-income securities, John Wiley & Sons. England.

Nomura Global Economics, Europe Will Work, Nomura, March 2011.

Norden, L. and M. Weber, 2009. The Co-movement of Credit Default Swap, Bond and Stock Markets: An Empirical Analysis, European Financial Management, Vol. 15, 477–485.

Pope, P. and P. Yadav, 1994. Discovering Errors in Tracking Error, Journal of Portfolio Management, Vol. 20, 27-32.

Potreba, M.J. and J.B. Shoven, 2002. Exchange Traded Funds: A New Investment Option for Flexible Investors, American Economic Review, 92, 422-427.

Rompotis, G., 2008. Performance and Trading Characteristics of German Passively Managed ETFs, International Research Journal of Finance and Economics, Issue 15, 210-223.

Schuknecht, L., von Hagen, J., Wolswijk G., 2010. Government bond risk premiums in the EU revisited. The impact of the financial crisis, ECB Working paper No. 1152 February.

The EDHEC European ETF Survey 2009. May 2009.

http://indices.markit.com/download/products/guides/Markit_iBoxx_EURLiquid_Guide.pdf

http://www.dax-indices.com/EN/MediaLibrary/Document/ebrexx_L_3_8_e.pdf

http://www.etftrends.com

Table 1 Summary of government bond indices together with index selection rules

This table presents a summary of index families together with index selection rules. F=France; G=Germany, I=Italy, N=Netherlands; S=Spain; Gr=Greece; Ir=Ireland; P=Portugal; Fin=Finland.

Index family	Prices			Min amount	Rating	Exposure (Maximum)	Country	Coupons' reinvestment
	Source	Adjustment	Type					
Barclays	Barclays capital	Daily	Mid	€2billion	Lower of S&P and Moody's	30% per issue	F, G, I, N, S	Monthly at rebalancing
Markit iBoxx Liq. EUR Sov.	Consortium	Per minute	Bid	€2billion	Lower of: S&P, Moody's, Fitch	20% per country; 1 bond per issuer; 3 bonds per country	Euro zone	Monthly at rebalancing
Markit iBoxx EUR Sovereign	Consortium	Per minute	Bid	€2billion	Lower of: S&P, Moody's, Fitch	-	Euro zone	Monthly at rebalancing
eb.rexx German Gov.	Eurex Bonds	Per minute	Best bid	€4billion	Lowest grade	30% per issue	G	Monthly at rebalancing
EuroMTS	MTS markets	Per 30 seconds	Best bid	€4billion	-	2 bonds per issuer	Euro zone	Overnight

Table 2 Aggregate country exposure of government bond indices tracked by sample ETFs, stratified by ETF providers

Snapshot calculated by authors; based on data provided in ETFs' prospectuses, as in May 2010.

	Germany	Italy	France	Spain	Belgium	Netherlands	Greece	Portugal	Austria	Ireland	Finland
iShares											
Barclays Term 1-3	32.6	42.1	11.1			14.2					
Barclays Term 3-5	32.2	23.7	30.8	13.3							
Barclays Term 5-7	48.0	25.0	12.7			14.2					
Barclays Term 7-10	56.6	7.9	29.9			5.6					
Barclays Term 10-15	9.3	41.4	30.3	7.0		12.0					
Barclays Term 15-30	29.4	34.8	21.7	9.7		4.4					
iBoxx€ Liq Sov Cap 1.5-2.5	20.3	20.0	20.2	8.0	4.4	9.5	13.9			3.8	
iBoxx€ Liq Sov Cap 2.5-5.5	20.3	20.0	20.5	20.0	13.0		6.1				
iBoxx€ Liq Sov Cap 5.5-10.5	20.5	20.3	20.3	19.9		13.1	5.9				
iBoxx€ Liq Sov Cap 10.5+	20.6	19.8	20.1	19.3	14.5	5.7					
iBoxx€ Liq Sov Cap 1.5-10.5	20.4	20.2	20.3	19.8		11.9	7.4				
eb.rexx 1.5-2.5	100.0										
eb.rexx 2.5-5.5	100.0										
eb.rexx 5.5-10.5	100.0										
eb.rexx 10.5+	100.0										
eb.rexx DE	100.0										
db-trackers											
Short iBoxx € Sov	21.6	23.7	20.9	9.4	5.9	5.4	3.9	2.2	3.7	2.0	1.1
iBoxx € Sov	21.6	23.7	20.9	9.4	5.9	5.4	3.9	2.2	3.7	2.0	1.1
iBoxx € Sov 1–3	25.1	24.4	19.8	11.1	5.7	5.5	3.8	1.8	1.2	1.0	0.7
iBoxx € Sov 3–5	24.4	17.2	19.8	11.0	7.2	5.4	4.2	2.3	4.3	2.0	
iBoxx € Sov 5–7	21.1	20.1	22.4	6.8	8.6	6.0	3.9	2.5	5.7	1.7	1.2
iBoxx € Sov 7–10	19.3	25.3	21.2	8.3	4.3	5.8	4.5	2.5	4.1	3.1	1.7
iBoxx € Sov 10–15	4.3	26.0	23.4	7.2	5.8	6.6	5.8	4.8	7.2	7.6	1.3
iBoxx € Sov 15+	23.9	30.1	21.3	9.4	4.6	3.7	2.4	0.9	3.4		0.5
iBoxx € Sov 25+	22.0	20.7	31.0	11.6		4.9	4.4	2.3	3.0		
Lyxor											
EuroMTS 1-3Y	23.7	23.4	21.8	11.5	6.2	4.3	4.6	0.9	1.3	1.84	0.7
EuroMTS 3-5Y	24.1	18.0	21.2	9.0	7.6	6.0	3.7	2.6	4.7	1.34	1.8
EuroMTS 5-7Y	20.8	16.9	23.6	6.1	8.9	6.2	5.5	2.8	6.2	1.91	1.2
EuroMTS 7-10Y	18.6	24.2	19.6	9.1	4.1	5.6	5.3	3.3	3.9	4.36	2.1
EuroMTS 10-15Y	2.1	31.0	23.2	5.2	6.8	8.5	5.3	4.6	8.4	5.0	
EuroMTS 15Y+	22.3	29.0	21.3	11.0	5.3	3.7	2.3	0.9	3.6		0.8

Table 3 Standard deviation (volatility) and correlation of sample ETFs and respective Indices, 2007 – 2010

This table presents volatility of sample ETFs and respective indices together with their correlations during the sample period. The volatility of ETF and respective indices is the annualized standard deviation of daily returns. Volatility calculations assume 252 trading days per annum. The correlation is calculated using daily time series of ETFs' NAVs.

ETFs and Indices	**Correlation**					**St.deviation**	
						ETF	**Index**
	2007	2008	2009	2010	2007-10	2007-10	2007-10
Ishares							
Barclays Term 1-3	0.99	0.99	1.00	1.00	0.99	1.43	1.42
Barclays Term 3-5	0.99	0.99	1.00	1.00	0.99	2.99	2.97
Barclays Term 5-7	**-**	**-**	**-**	0.99	0.97	3.27	3.29
Barclays Term 7-10	0.99	0.99	1.00	0.99	0.99	5.19	5.20
Barclays Term 10-15	-	-	-	1.00	0.98	4.71	4.70
Barclays Term 15-30	0.99	0.99	1.00	1.00	0.99	8.93	8.95
Average for Barclays	**0.99**	**0.99**	**1.00**	**0.99**	**0.99**	**4.42**	**4.42**
iBoxx€ Liq Sov Cap 1.5-2.5	0.95	0.98	0.98	1.00	0.99	2.34	2.42
iBoxx€ Liq Sov Cap 2.5-5.5	0.99	0.99	0.98	0.98	0.99	2.94	2.97
iBoxx€ Liq Sov Cap 5.5-	0.92	0.99	0.96	1.00	0.97	4.79	4.98
iBoxx€ Liq Sov Cap 10.5+	0.84	0.99	0.93	1.00	0.94	8.73	9.32
iBoxx€ Liq Sov Cap 1.5-	0.93	0.99	0.96	1.00	0.98	4.33	4.50
Average for iBoxx	**0.92**	**0.99**	**0.96**	**0.99**	**0.97**	**4.63**	**4.84**
eb.rexx 1.5-2.5	0.97	0.98	0.98	0.98	0.98	1.51	1.54
eb.rexx 2.5-5.5	0.99	0.99	0.98	1.00	0.99	2.92	3.01
eb.rexx 5.5-10.5	1.00	0.99	0.96	1.00	0.98	4.94	5.08
eb.rexx 10.5+	1.00	0.96	0.92	1.00	0.96	10.02	10.31
eb.rexx DE	0.99	0.99	0.98	1.00	0.99	3.36	3.40
Average for eb.rexx	**0.99**	**0.98**	**0.96**	**0.99**	**0.98**	**4.55**	**4.67**
db x-trackers							
short iBoxx	-	0.97	0.98	0.99	0.98	4.40	4.40
iBoxx € Sov	0.99	0.98	0.98	0.99	**0.98**	4.25	4.23
iBoxx € Sov 1–3	0.96	0.98	0.98	0.99	0.98	1.68	1.67
iBoxx € Sov 3–5	0.98	0.97	0.98	0.99	0.98	3.28	3.25
iBoxx € Sov 5–7	0.97	0.95	0.98	1.00	0.96	4.38	4.34
iBoxx € Sov 7–10	0.99	0.98	0.99	1.00	0.98	5.29	5.27
iBoxx € Sov 10–15	0.99	0.98	0.99	1.00	0.99	6.46	6.44
iBoxx € Sov 15+	0.99	0.99	0.98	0.98	0.99	9.21	9.20
iBoxx € Sov 25+	1.00	0.98	0.98	0.98	0.98	11.08	11.02
Average for db x	**0.98**	**0.98**	**0.98**	**0.99**	**0.98**	**5.56**	**5.54**
Lyxor							
EuroMTS 1-3	0.46	0.46	0.47	0.27	0.43	1.66	1.58
EuroMTS 3-5	0.43	0.47	0.38	0.34	0.42	3.14	2.97
EuroMTS 5-7	0.43	0.48	0.40	0.28	0.42	4.28	4.12
EuroMTS 7-10	0.40	0.44	0.43	0.39	0.43	5.08	4.90
EuroMTS 10-15	0.39	0.39	0.44	0.45	0.41	6.24	6.07
EuroMTS 15+	0.74	0.98	0.99	0.96	0.96	10.31	10.28
Average for Lyxor	**0.48**	**0.54**	**0.52**	**0.45**	**0.51**	**5.12**	**4.99**

Figure 1 Changes in volatility for sample ETFs
This figure represents changes in volatility for the sample ETFs targeting 10+ year maturities. The volatility is estimated as a rolling 3-month standard deviation of daily returns (annualized).

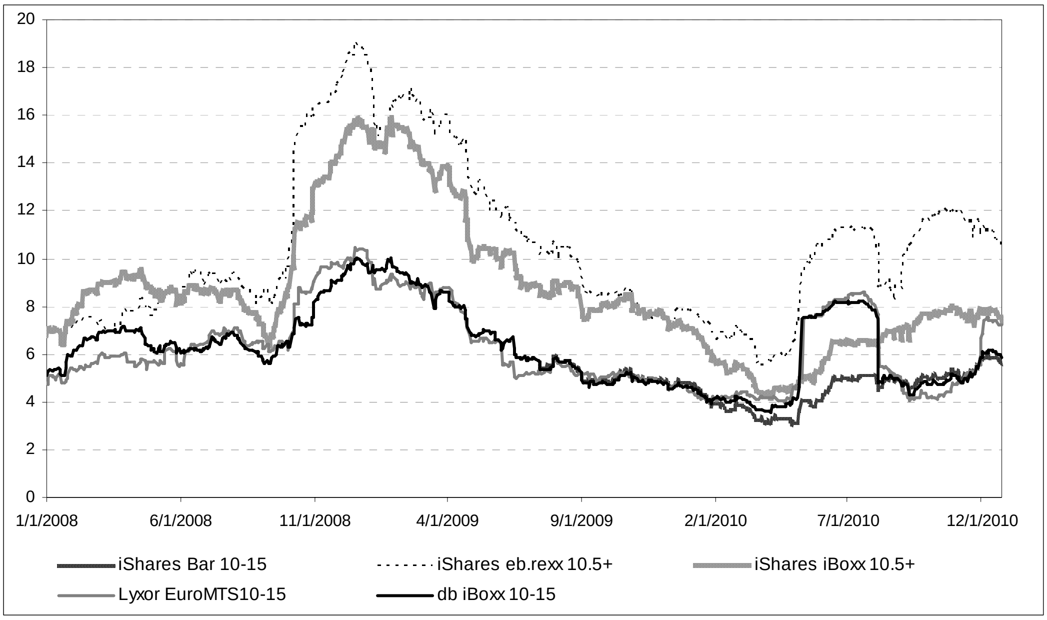

Table 4 Performance and sample TE1, 2007 – 2010
This table represents ETFs annual returns (%) and TE1 presented in basis points [in brackets]. Positive basis points indicate undeperformance of the funds over underlying indices.

	Year	iShares			db x-track	Lyxor
		Barclays T.	iBoxx L.S.C.	eb.rexx		
Mean	2007	1.10 [23]	1.35 [8]	1.51 [5]	5.00 [16]	2.29 [-7]
	2008	9.99 [28]	8.99 [24]	12.35 [32]	8.17 [17]	9.13 [17]
	2009	4.39 [28]	3.64 [9]	1.91 [7]	2.79 [15]	4.46 [3]
	2010	12.45 [20]	7.88 [16]	14.58 [68]	7.80 [15]	8.37 [33]
	2007-10	6.58 [27]	4.95 [13]	6.10 [19]	5.37 [16]	5.86 [10]
Median	2007	2.15 [23]	1.89 [7]	2.66 [9]	5.03 [16]	2.97 [10]
	2008	10.37 [29]	9.77 [24]	11.52 [36]	8.97 [17]	9.43 [17]
	2009	4.21[27]	3.73 [8]	2.78 [10]	4.19 [15]	4.81 [0]
	2010	14.08 [21]	6.53 [14]	12.03 [63]	9.37 [16]	8.12 [31]
	2007-10	6.51 [26]	5.22[13]	6.29[16]	6.62 [16]	5.91[12]
Min	2007	-3.42 [19]	-3.29 [3]	-3.59 [-11]	3.42 [15]	-0.2 [99]
	2008	6.42 [24]	6.48 [18]	7.13 [15]	-3.54 [12]	6.48 [15]
	2009	2.23 [21]	1.19 [6]	-1.59 [-12]	-3.21 [13]	2.13 [-7]
	2010	4.14 [13]	1.87 [10]	5.42 [29]	-7.77[12]	3.36 [4]
Max	2007	3.52 [26]	3.42 [12]	3.51 [13]	6.44 [16]	3.55 [20]
	2008	12.8 [31]	11.53 [27]	18.44 [47]	13.81 [19]	10.76 [18]
	2009	7.39 [39]	5.34 [11]	3.47 [15]	5.29 [16]	5.40 [17]
	2010	16.38 [25]	17.21 [30]	28.02 [127]	14.37 [18]	13.51 [63]

Table 5 Sample TE2, 2007 – 2010

This table presents results for average (mean) 3 month TE2 for respective ETFs, based on monthly and daily NAV series. N is the number of bonds in the portfolio – calculated by authors based on the data from ETFs' prospectuses as in May 2010. P-values for one sample T-test for mean=0 vs. mean#0 in brackets. Unreported results for one sample Wilcoxon test for median=0 vs. median#0 are economically and statistically consistent with the reported results for the T-test. *indicate significance at the 10% level;** indicate significance at the 5% level; *** indicate significance at the 1% level.

		AVERAGE 3-MONTH TE2 (IN BPS)					
	N	TE2 (monthly)	TE2 (daily)				
		2007-10	2007-10	2007	2008	2009	2010
iShares							
Barclays Term 1-3	10	1.23***	0.77***	0.69	1.37	0.44	0.2
Barclays Term 3-5	15	1.96***	1.53***	1.28	3.02	0.59	0.54
Barclays Term 5-7	9	2.38***	2.29***	-	-	3.48	0.83
Barclays Term 7-10	13	2.80***	2.33***	2.08	5.08	0.36	0.76
Barclays Term 10-15	14	3.32***	2.95***	-	-	4.43	1.12
Barclays Term 15-30	30	4.69***	4.08***	4.14	8.77	0.62	0.68
Average for Barclays	**15**	**2.73*** **	**2.32*** **	**2.05*** **	**4.56*** **	**1.65*** **	**0.69**** **
iBoxx€ Liq Sov Cap 1.5-	15	1.97***	1.87***	1.84	2.28	1.55	1.72
iBoxx€ Liq Sov Cap 2.5-	15	2.63***	2.52***	1.81	2.81	2.93	2.11
iBoxx€ Liq Sov Cap 5.5-	15	7.05***	5.30***	7.53	4.26	6.02	1.75
iBoxx€ Liq Sov Cap	15	17.46***	12.80***	19.65	7.9	16.61	2.1
iBoxx€ Liq Sov Cap 1.5-	25	6.03***	4.65***	6.20	3.96	5.29	1.72
Average for iBoxx	**17**	**7.03*** **	**5.43*** **	**7.40*** **	**4.24*** **	**6.48*** **	**1.88*** **
eb.rexx 1.5-2.5	6	1.77***	1.89***	1.55	2.33	1.92	1.33
eb.rexx 2.5-5.5	12	3.15***	2.57***	1.74	2.63	3.56	1.47
eb.rexx 5.5-10.5	10	6.54***	4.84***	1.68	4.11	9.18	1.62
eb.rexx 10.5+	10	15.66***	11.78***	2.23	14.23	20.09	2.38
eb.rexx DE	25	2.77***	2.90***	1.54	3.09	4.28	1.44
Average for eb.rexx	**13**	**5.98*** **	**4.79*** **	**1.75*** **	**5.28*** **	**7.8*** **	**1.65*** **
db x-trackers							
short iBoxx	252	4.54***	4.38***	-	6.81	4.81	0.56
iBoxx € Sov	252	3.49***	3.32***	3.69	4.85	2.72	0.48
iBoxx € Sov 1–3	60	1.57***	1.53***	2.52	2.05	1.01	0.6
iBoxx € Sov 3–5	48	3.03***	3.07***	4.06	4.67	2.03	0.65
iBoxx € Sov 5–7	33	5.34***	5.03***	5.26	9.03	2.71	0.39
iBoxx € Sov 7–10	47	4.10***	3.81***	3.89	6.2	2.63	0.55
iBoxx € Sov 10–15	24	4.55***	4.34***	3.88	6.74	3.55	0.51
iBoxx € Sov 15+	44	6.33***	6.40***	5.45	7.69	7.56	0.79
iBoxx € Sov 25+	18	11.74***	9.42***	5.90	13.86	9.2	1.3
Average for db iBoxx	**86**	**4.97*** **	**4.59*** **	**4.33*** **	**6.88*** **	**4.02*** **	**0.65*** **
Lyxor							
EuroMTS 1-3	20	10.08****	9.54***	6.97	12.76	9.07	7.37
EuroMTS 3-5	22	20.33***	19.34***	13.68	25.15	20.07	13.27
EuroMTS 5-7	20	28.61***	26.94***	20.60	32.61	28.39	19.82
EuroMTS 7-10	22	34.78***	32.65***	26.20	38.86	34.22	24.02
EuroMTS 10-15	18	44.40***	40.72***	31.52	49.46	43.12	29.6
EuroMTS 15+	36	15.77***	12.87***	34.85	15.88	5.68	8.09
Average for Lyxor	**23**	**25.66*** **	**23.68*** **	**22.30*** **	**29.12*** **	**23.42*** **	**17.03*** **

Table 6 Differences in average TE2 across index families

This table presents results for the two sample T test for difference in mean and Mann Whitney (MW) test for difference in median TEs for different families of ETFs. All tests are for TEs estimated based on daily NAV series. *indicate significance at the 10% level;** indicate significance at the 5% level; *** indicate significance at the 1% level.

		T-test	MW test
Barclays T. vs.	iBoxx Liq.Sov.	-13.50***	16.31***
	eb.rexx	-14.26***	15.92***
	db.iBoxx	-11.29***	2.33**
	EuroMTS	-110.00***	33.55***
eb.rexx vs.	iBoxx Liq.Sov.	-2.57***	-3.93***
	db.iBoxx	-0.74	-5.29***
	EuroMTS	-80.40***	33.60***
iBoxx Liq.Sov vs.	Db.iBoxx	3.01***	-7.19***
	EuroMTS	-69.29***	32.29***
db.iBoxx vs.	EuroMTS	-74.06***	32.28***

Figure 2 Patterns of sample TE2, 2007-2010
This figure presents the patterns of average three month TE2 (in bps) for the sample ETFs targeting 10+ year maturities.

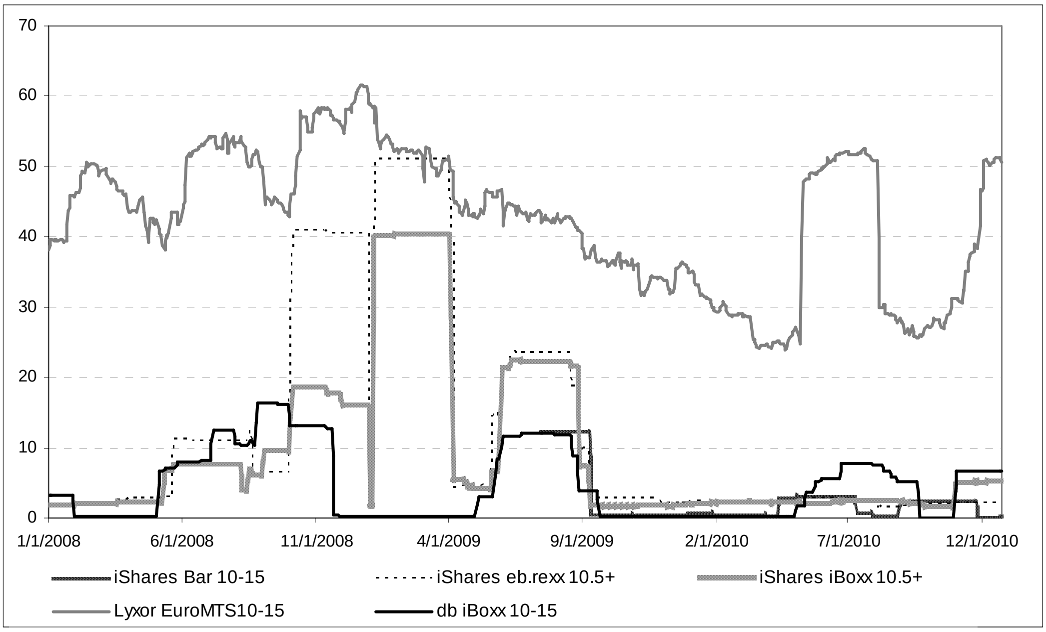

Table 7 OLS Regression model for ETFs' returns and TE3, 2007 – 2010

This table presents results of the OLS model for ETFs daily returns as dependent variable and daily returns of the underlying indices as explanatory variable: $r_p = \alpha_i + \beta_i r_b + \varepsilon_p$. TE3 is standard error of regression (i.e. standard deviation of residuals) in basic points. T-tests are two tail tests for: i) Alpha =0 vs. Alpha≠0 (column 3); ii) beta=1 vs. beta ≠1 (column 5). *indicate significance at the 10% level;** indicate significance at the 5% level; *** indicate significance at the 1% level.

	Alpha	T-test	Beta	T-test	R^2	TE3	N
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
iShares							
Barclays Term 1-3	0.0000	-2.56**	1.00	-1.15	0.99	1.01	1037
Barclays Term 3-5	0.0000	-0.93	1.00	-0.55	0.98	2.32	1037
Barclays Term 5-7	0.0000	-0.22	0.98	-2.33**	0.96	4.23	439
Barclays Term 7-10	0.0000	-0.54	0.99	-2.24**	0.98	4.07	1037
Barclays Term 10-15	0.0000	-0.35	0.99	-1.33	0.98	4.84	439
Barclays Term 15-30	0.0000	-0.41	0.99	-2.37**	0.99	5.78	1037
Average	0.0000	-12.70***	0.99	-3.18**	0.98	1.01	
iBoxx							
iBoxx€ Liq Sov Cap 1.5-2.5	0.0000	0.20	0.96	-9.82***	0.98	1.99	1037
iBoxx€ Liq Sov Cap 2.5-5.5	0.0000	-0.13	0.98	-4.80***	0.97	2.91	1037
iBoxx€ Liq Sov Cap 5.5-10.5	0.0000	0.22	0.94	-8.83***	0.95	6.64	1037
iBoxx€ Liq Sov Cap 10.5+	0.0000	0.17	0.89	-11.50***	0.89	17.52	1037
iBoxx€ Liq Sov Cap 1.5-10.5	0.0000	0.19	0.94	-8.89**	0.96	5.58	1037
Average	0.0000	1.97	0.94	-4.10**	0.95	6.93	
Eb							
eb.rexx 1.5-2.5	0.0000	0.10	0.96	-7.00***	0.96	1.80	1037
eb.rexx 2.5-5.5	0.0000	0.00	0.96	-9.26***	0.98	2.55	1037
eb.rexx 5.5-10.5	0.0000	0.10	0.96	-8.19***	0.97	5.47	1037
eb.rexx 10.5+	0.0000	0.17	0.94	-7.68***	0.93	16.88	1037
eb.rexx DE	0.0000	-0.03	0.98	-5.54***	0.98	3.06	1037
Average	0.0000	1.29	0.96	-7.20***	0.96	5.95	
db x-trackers							
short iBoxx	0.0000	-0.27	0.98	-2.85***	0.96	5.43	687
iBoxx € Sov	0.0000	-0.24	0.99	-2.22**	0.97	4.64	937
iBoxx € Sov 1–3	0.0000	-0.45	0.98	-2.57**	0.96	2.04	934
iBoxx € Sov 3–5	0.0000	-0.18	0.98	-2.32**	0.95	4.22	934
iBoxx € Sov 5–7	0.0000	0.00	0.97	-3.24***	0.93	6.79	931
iBoxx € Sov 7–10	0.0000	-0.22	0.99	-1.97**	0.97	5.49	931
iBoxx € Sov 10–15	0.0000	-0.21	0.99	-2.13**	0.97	6.35	929
iBoxx € Sov 15+	0.0000	-0.10	0.99	-2.50**	0.97	9.87	929
iBoxx € Sov 25+	0.0000	-0.07	0.98	-2.35**	0.96	13.75	927
Average	0.0000	-6.37***	0.98	-9.21***	0.96	6.51	
Lyxor							
EuroMTS 1-3	0.0001	2.65**	0.44	-18.98***	0.17	9.21	1037
EuroMTS 3-5	0.0001	1.72*	0.44	-18.95***	0.18	17.34	1037
EuroMTS 5-7	0.0001	1.36	0.44	-19.21***	0.18	23.76	1023
EuroMTS 7-10	0.0001	1.01	0.44	-19.34***	0.18	28.23	1023
EuroMTS 10-15	0.0001	0.62	0.43	-19.51***	0.18	35.58	1037
EuroMTS 15+	0.0000	0.13	0.97	-3.75***	0.93	16.91	911
Average	0.0001	5.23***	0.53	-5.38***	0.30	21.84	

Table 8 Cointegrated vector error-correction (VEC) model and TE4

This table presents results of Johansen test (JT) for cointegration between daily log NAV values for ETFs and relevant indices together with the results of the VEC model. Estimations are made assuming a constant trend. Critical value for JT trace statistics for rejecting a null (at 5% level of significance) that there is zero cointegrating vector is 15.41. Number of lags (NL) is determined based on unreported results of Hannan-Quin information criterion (HQIC), Schwartz Bayesian information criterion (SBIC) and sequential Likelihood Ratio (LR) methods. Log likelihood is for the overall model. Estimated parameters of the cointegrating vector are normalised cointegrating vector coefficients for ETFs and indices. The coefficient on ETF is restricted to unity by the Johansen normalization. DF is Dickey-Fuller test statistics from unit root tests for the cointegrating regressions' residuals (TE4); *indicate significance at the 10% level;** indicate significance at the 5% level; *** indicate significance at the 1% level.

	NL	JT trace	Normalised coefficients		DF	Log lik.	N
			ETF	Index			
iShares							
Barclays Term 1-3	4	5.79	-	-	-	-	1037
Barclays Term 3-5	4	12.41	-	-	-	-	1037
Barclays Term 5-7	4	24.46**	1	-0.9609***	-30.033	4,954.33	439
Barclays Term 7-10	4	13.67	-	-	-	-	1037
Barclays Term 10-15	4	16.23**	1	-0.9718***	-34.067	4,714.59	439
Barclays Term 15-30	4	14.94	-	-	-	-	1037
Average (mean)		**14.583**			**-32.050**		
iBoxx							
iBoxx€ Liq Sov Cap 1.5-2.5	4	19.82**	1	-0.9755***	-25.255	12,752.1	1037
iBoxx€ Liq Sov Cap 2.5-5.5	4	31.24**	1	-0.9740***	-20.314	12,169.1	1037
iBoxx€ Liq Sov Cap 5.5-10.5	4	125.03**	1	-0.9793***	-30.453	10,756.8	1037
iBoxx€ Liq Sov Cap 10.5+	4	151.25**	1	-0.9820***	-29.650	9,124.92	1037
iBoxx€ Liq Sov Cap 1.5-10.5	4	151.22**	1	-1.0574***	-29.606	12,414.8	1037
Average (mean)		**95.712**			**-27.056**		
Eb							
eb.rexx 1.5-2.5	4	17.88**	1	-0.9655***	-29.716	13,292.1	1037
eb.rexx 2.5-5.5	4	27.67**	1	-0.9626***	-32.244	12,247.5	1037
eb.rexx 5.5-10.5	4	127.01**	1	-0.9660***	-24.649	10,997.6	1037
eb.rexx 10.5+	4	172.03**	1	-0.9765***	-29.972	9,078.45	1037
eb.rexx DE	4	29.41**	1	-0.9774***	-29.979	11,952.7	1037
Average (mean)		**74.800**			**-29.312**		
db x-trackers							
short iBoxx	4	46.07**	1	-1.0273***	-30.982	7,427.96	687
iBoxx € Sov	4	35.27**	1	-0.9776***	-32.723	10,271.4	937
iBoxx € Sov 1–3	4	9.42	-	-	-	-	934
iBoxx € Sov 3–5	4	25.56**	1	-0.9781***	-35.604	10,582.2	934
iBoxx € Sov 5–7	4	56.85**	1	-0.9802***	-36.015	9,869.88	931
iBoxx € Sov 7–10	4	36.93**	1	-0.9796***	-33.610	9,855.68	931
iBoxx € Sov 10–15	4	38.78**	1	-0.9782***	-31.539	9,519.44	929
iBoxx € Sov 15+	4	80.99**	1	-0.9809***	-27.250	8,806.23	929
iBoxx € Sov 25+	4	103.73**	1	-0.9808***	-24.073	8,335.31	927
Average (mean)		**48.178**			**-30.497**		
Lyxor							
EuroMTS 1-3	4	75.79**	1	-0.9664***	-30.409	11,987.8	1037
EuroMTS 3-5	4	169.30**	1	0.9731***	-31.854	10,723.3	1037
EuroMTS 5-7	3	283.29**	1	-0.9786***	-29.626	9,933.10	1023
EuroMTS 7-10	3	278.16**	1	-0.9774***	-29.696	9,575.01	1023
EuroMTS 10-15	3	261.71**	1	-0.9723***	-29.634	9,241.08	1037
EuroMTS 15+	1	360.42**	1	-0.9792***	-22.996	8,016.32	911
Average (mean)		**238.12**			**-28.761**		

Table 9 Correlation matrix of variables used for Panel regressions, 2008-2010
This table presents the correlation matrix for all variables used in the Panel data regressions.

	LNWCD	BID-ASK	LNTERM	SYNTHETIC	GCRISIS	LCRISIS	WINDEX	EB.REXX	WPCDS
LNWCD	1								
BID-ASK	0.1506	1							
LNTERM	0.0820	0.2228	1						
SYNTHETIC	0.2186	0.0896	0.1192	1					
GCRISIS	**0.4737**	0.0765	0.0187	-0.0222	1				
LCRISIS	0.1250	0.0208	-0.0185	0.0157	-0.04238	1			
WINDEX	0.0075	0.0422	0.0182	0.0049	-0.0528	0.0730	1		
EB.REXX	**-0.4636**	-0.0765	-0.1064	-0.3614	-0.0196	0.0139	-0.0174	1	
WPCDS	-0.0309	-0.0040	-0.0045	-0.0143	-0.0126	-0.0175	0.0367	0.0422	1

Table 10 Panel data analysis of sample tracking errors (TE2 and TE4), 2008- 2010

This table presents results of the random effects GLS regression with a robust estimator of variance adjusted for clustering on fund. Unreported results for Hausman test for systematic difference in coefficients (Ho=difference in coefficients is not systematic), and Breusch and Pagan Langrangian (BPL) multiplier test for random effects support the choice of the random (over fixed) effects model. TE2 is based on daily returns and is a dependent variable in all models. TE4 are residuals from the cointegrating regression; Panel A presents results for models with TE2 as a dependent variable;Panel B presents results for models with TE4 as a dependent variable; GCRISIS is a categorical variable equal to 1 for observations during the post-Greek sovereign crisis (19th April 2010 to 23rd December 2010), and 0 otherwise; LCRISISis a categorical variable equal to 1 for observations during the period of credit market crisis (1st September 2008 to 30th September 2009), and 0 otherwise; WINDEX is the weekly percentage change of the relevant index in the previous week; EB.REXX is a categorical equal 1 if ETF is from eb.rexx family, and 0 otherwise; LNWCDS is the natural logarithm of the weighted average CDS spreads for ETFs; WPCDS is the weekly percentage change in the level of the weighted average CDS spreads; WINDEX*LNWCDS is an interaction variable constructed from WINDEX and LNWCDS; *indicate significance at the 10% level; ** indicate significance at the 5% level; *** indicate significance at the 1% level.

	Panel A				
	TE2	TE2	TE2	TE2	TE2
BID-ASK	0.000045**	0.000030**	0.000041	0.000037	0.000026
LNTERM	0.00033	0.000343	0.000452	0.000471	0.000467
SYNTHETIC	0.000833**	0.000810**	0.002107***	0.002089***	0.002077***
GCRISIS		0.000022**		0.000104***	-0.000032***
LCRISIS		0.000408***		0.000490***	0.000413***
LNWCDS			-0.000027***	-0.000111***	
WPCDS			-0.000008	0.000001	-0.000003
WINDEX	0.004293***	0.002699***	0.004545***	0.002640**	0.005885
WINDEX*LNWCDS					-0.000854
EB.REXX			0.000245	0.000148	0.000242
Constant	-0.000793	-0.00098	-0.001191	-0.001091	-0.001466
Wald Chi2(4)	40.51***	2,710.75***	68.04***	3,304.78***	2,624.44***
R^2	0.176	0.2021	0.5499	0.5750	0.5752
N (obs.)	22,825	22,825	15,973	15,973	15,973
N (groups)	31	31	22	22	22

	Panel B				
	TE4	TE4	TE4	TE4	TE4
BID-ASK	-0.000061***	-0.000059***	-0.000076**	-0.000081**	-0.000091**
LNTERM	0.000083***	0.000081***	0.000057***	0.000063***	0.000058***
SYNTHETIC	0.000045*	0.000054***	0.000091***	0.000097***	0.000093***
GCRISIS		0.000057**		0.000201***	0.000078***
LCRISIS		-0.000029		0.000094***	0.000023
LNWCDS			-0.000032*	-0.000010***	
WPCDS			-0.000005	-0.000003	-0.000004
WINDEX	0.015091***	0.015319***	0.025635***	0.025792***	-0.000274
WINDEX*LNWCDS					0.006894***
EB.REXX			0.000037**	-0.000025	0.000067***
Constant	-0.000119*	-0.000114	0.000068	0.000247***	-0.000093*
Wald Chi2(4)	138.25***	159.71***	265.39***	321.35***	297.11***
R^2	0.0139	0.0145	0.0276	0.0297	0.0291
N (obs.)	17,231	17,231	12,112	12,112	12,112
N (groups)	24	24	16	16	16