APPLIED FINANCE LETTERS

VOLUME 13, 2024

PERFORMANCE AND TRACKING EFFICIENCY OF COMMODITY ETFS IN THE UK

GERASIMOS G. ROMPOTIS[1*]

1. National and Kapodistrian University of Athens, Greece.

* Corresponding Author: Gerasimos G. Rompotis, Department of Economics, National and Kapodistrian University of Athens: Sofokleous Str (10559), Greece.
✉ geras3238@yahoo.gr

Abstract

This paper examines the performance and tracking efficiency of twenty eight iShares ETFs traded on the London Stock Exchange in the UK. The results indicate that, on average, the performance of the examined ETFs has been positive during their entire trading history. However, these ETFs have failed to fully replicate the performance of the underlying commodities and indexes. At the cumulative level, an average underperformance of 320 basis points is found. In addition, at the sample level, about 52% of daily tracking errors are negative (indicating underperformance), and 47% of tracking errors are positive (reflecting outperformance). Based on our results, the tracking error is induced by the departure from the full replication of the underlying assets. In addition, tracking error is found to be positively related to the age of ETFs but negatively to their assets. It is also found that ETFs applying physical replication have relatively lower tracking errors than ETFs pursuing synthetic replication. Finally, no significant differences are found in tracking errors among the managing companies of commodity ETFs.

Keywords: ETFs, commodities, performance, tracking error

1. Introduction

This study focuses on Exchange Traded Funds (ETFs) which invest in commodities. Investors use commodities tools to diversify their portfolios. In addition, during highly volatile equity markets, investing in commodities can act as a relatively safe haven, even though commodities themselves are not risk-free investments. The prices of commodities are affected by several factors, such as unusual weather conditions, natural disasters, unsuitable agricultural techniques, pollution, human activity, political and economic crises, and war conflicts.

Publicly traded commodities include metals, energy, livestock, meat and agricultural products. Access to commodity markets is attained in several ways including the physical purchase of a commodity, as well as investing in futures contracts, options and commodity ETFs. A commodity ETF invests in agricultural products, natural resources and metals. The key benefits of commodity ETFs concern the potential for high portfolio diversification, low cost, variety of assets, continuous trading, high liquidity and tax efficiency.

Commodity ETFs attain exposure to the desired commodities either by physically storing the selected commodity, or via investing in futures contracts. The latter is the most commonly adopted option among commodity ETFs and has the benefit of avoiding the storage costs regarding the physical exposure. However, this "futures-based" approach is subject to the "rolling costs" relating to rolling

the expiring futures contracts by closing them out and reopening them as future dated ones. Finally, several ETFs choose to get access to commodities by tracking relevant commodity indexes.

The performance of commodity ETFs can be affected by several factors. The difference between the spot and future prices of the underlying commodities is one of these factors. Money market (collateral) yield and the rolling yield also affect ETFs' performance. Money market yield is the revenue gained via investing the underlying assets of a commodity ETF in interest bearing accounts, including Treasury Bills or Treasury Inflation-Protected Securities (TIPS). Rolling yield refers to the gains and losses from rolling the expiring futures contracts. Developments in equity markets can also bear an impact on the performance of commodity ETFs.

The performance of commodity ETFs has been evaluated by several studies. Sousa (2014) shows that metal ETFs traded on the NYSE Arca have negative but statistically insignificant alphas, while their tracking error is low. Neff and Isengildina-Massa (2018) also find that the average tracking error of commodity ETFs is small even though, occasionally, tracking errors can be quite large. Rompotis (2016) reveals that the physically backed commodity ETFs perform better than their futures-based peers. He also finds that the tracking error of futures-based ETFs is significantly higher than that of the physically backed ETFs. Similar results are reported by Fassas (2014). On the UK-listed equity ETFs, by investigating the tracking performance of physical and synthetic ETFs during the period 2008-2013, Mateus and Rahmani (2015) find no significant differences in their ability to replicate the performance of their benchmarks. Similar results are provided by Maurer and Williams (2015). Merz (2015) investigates the tracking risk of physical and synthetic European ETFs and provides evidence that ETFs that follow a synthetic replication strategy, rather than holding the underlying securities comprising the benchmark, are less prone to tracking error. However, in most cases, they underperform both the benchmarks and their physical counterparts.

Furthermore, Perera et al. (2022) note that the replication method, along with the volatility in the prices of the underlying commodities, can affect the tracking ability of agricultural ETFs. They also show that the tracking error of these ETFs is not trivial, but it does not last very long. Stewart et al. (2023) show that the tracking error of commodity ETFs and Exchange Traded Notes (ETNs) focusing on the agricultural and energy sectors is driven by the arbitrage process inherent to these products. The authors also report no material differences in the tracking ability across agricultural and energy ETFs. Guo and Leung (2015) show that the leveraged commodity ETFs underperform their benchmarks in the long run. Similar results are reported by Murphy and Wright (2010). In this respect, Guedj et al. (2011) note that it is not easy for futures-based commodity ETFs to replicate their benchmarks in the long run because the term structure of futures contracts may lead to large deviations between the price of ETFs and the spot price of the underlying commodities.

In this paper, we examine the performance and tracking efficiency of twenty eight ETFs that are traded on the London Stock Exchange (LSE). These ETFs are the so-called "iShares", which are managed by BlackRock. To the best of our knowledge, this is the first study on the commodity ETFs listed in the UK. The results show that the examined ETFs fail to fully replicate the performance of the underlying commodities and indexes. The average cumulative underperformance equals 320 basis points (bps). Underperformance is also verified by the fact that the number of days with negative raw tracking errors is on average higher than the number of days with positive raw tracking errors.

2. Data and Methodology

2.1 Data and Descriptive Statistics

The sample of our study includes twenty eight commodity iShares traded on the LSE.[1] Table 1 presents the profiles of these ETFs. Twenty ETFs are physically exposed to precious metals, including gold, silver, platinum and palladium. Three futures-based ETFs track relevant commodity indexes, while five synthetic ETFs invest in commodities including cotton, copper, coffee, sugar and crude oil. The oldest ETF in the sample is about 19.8 years old, while the newest one is just 1.5 years old. Moreover, the largest ETF is the Invesco Physical Gold ETC, whose assets on 31 December 2023 amounted to $14.2 billion. The average ETF in the sample held $2 billion on the same date. Finally, the average expense ratio of the examined ETFs is 0.30%, with minimum and maximum expense ratios being equal to 0.11% and 0.49%, respectively. Table 1 also reports the managing company of each ETF in the sample. One ETF is provided by HANETF. Four ETFs are managed by DWS. Invesco offers three ETFs. Nine ETFs (iShares) are provided by BlackRock. Finally, Wisdom Tree adds eleven ETFs to our sample.

Table 1: Profiles of ETFs

Symbol	Name	Benchmark	Provider	Replication	Inception Date	Age[1]	Assets ($M)[1]	Expense Ratio
RMAU	The Royal Mint Responsibly Sourced Physical Gold ETC	Gold Spot	HANETF	Physical	14/2/2020	3.88	692.89	0.25
XGLD	Xtrackers Physical Gold ETC	Gold Spot	DWS	Physical	15/6/2010	13.55	1,920.00	0.25
XGDU	Xtrackers IE Physical Gold ETC Securities	Gold Spot	DWS	Physical	22/4/2020	3.69	3,070.00	0.11
XPPT	Xtrackers IE Physical Platinum ETC Securities	Platinum	DWS	Physical	16/4/2020	3.71	14.28	0.38
XSLR	Xtrackers IE Physical Silver ETC Securities	Silver	DWS	Physical	29/4/2020	3.67	104.93	0.20
SGLD	Invesco Physical Gold ETC	Gold Spot	INVESCO	Physical	24/6/2009	14.53	14,200.00	0.12
SPPT	Invesco Physical Platinum ETC	Platinum	INVESCO	Physical	15/4/2011	12.72	21.02	0.19
SSLV	Invesco Physical Silver ETC	Silver	INVESCO	Physical	15/4/2011	12.72	161.60	0.19
IGLN	iShares Physical Gold ETC	Gold Spot	ISHARES	Physical	8/4/2011	12.74	13,050.00	0.12
IGLG	iShares Physical Gold GBP Hedged ETC	Gold Spot	ISHARES	Physical	5/7/2022	1.49	11.53	0.25
IPDM	iShares Physical Palladium ETC	Palladium	ISHARES	Physical	8/4/2011	12.74	15.61	0.20
IPLT	iShares Physical Platinum ETC	Platinum	ISHARES	Physical	8/4/2011	12.74	70.55	0.20
ISLN	iShares Physical Silver ETC	Silver	ISHARES	Physical	8/4/2011	12.74	515.09	0.20
ICOM	iShares Diversified Commodity Swap UCITS ETF	Bloomberg Commodity TRI	ISHARES	Synthetic	18/7/2017	6.46	1,330.00	0.19
ROLL	iShares Bl. Enh. Roll Yield Com. Swap UCITS ETF	Bloomberg Enhanced Roll Yield TRI	ISHARES	Synthetic	28/9/2018	5.26	1,240.00	0.28
EXXY	iShares Diversified Commodity Swap UCITS ETF (DE)	Bloomberg Commodity TRI	ISHARES	Synthetic	7/8/2007	16.41	256.09	0.46
IGLD	iShares Physical Gold EUR Hedged ETC	ICE LBMA Gold EUR Hedged Index	ISHARES	Physical	5/7/2022	1.49	24.64	0.25
GBS	Gold Bullion Securities	Gold Spot	WISDOM TREE	Physical	31/3/2004	19.76	2,630.00	0.40
PHAG	WisdomTree Physical Silver	Silver	WISDOM TREE	Physical	24/4/2007	16.70	1,150.00	0.49
PHAU	WisdomTree Physical Gold	Gold Spot	WISDOM TREE	Physical	24/4/2007	16.70	4,260.00	0.39
PHPD	WisdomTree Physical Palladium	Palladium	WISDOM TREE	Physical	24/4/2007	16.70	84.69	0.49
PHPT	WisdomTree Physical Platinum	Platinum	WISDOM TREE	Physical	24/4/2007	16.70	8,920.00	0.49
WGLD	WisdomTree Core Physical Gold	Gold Spot	WISDOM TREE	Physical	3/12/2020	3.08	625.72	0.12

[1] About 270 commodity ETFs (ETCs) are traded on the LSE. However, there are no publicly available data for the majority of these ETFs and especially for their benchmarks. As a corollary, our sample is a relatively small portion of the entire population of the UK-listed commodity ETFs.

Symbol	Name	Benchmark	Provider	Replication	Inception Date	Age[1]	Assets ($M)[1]	Expense Ratio
COTN	WisdomTree Cotton	Cotton	WISDOM TREE	Synthetic	27/9/2006	17.27	5.36	0.49
COPA	WisdomTree Coper	Copper	WISDOM TREE	Synthetic	27/9/2006	17.27	1,460.00	0.49
COFF	WisdomTree Coffee	Coffee	WISDOM TREE	Synthetic	27/9/2006	17.27	29.32	0.49
SUGA	WisdomTree Sugar	Sugar	WISDOM TREE	Synthetic	27/9/2006	17.27	10.04	0.49
BRND	WisdomTree Bloomberg Brent Crude Oil	Brent Crude Oil	WISDOM TREE	Synthetic	9/4/2015	8.73	10.67	0.25
Average						**11.36**	**1,995.86**	**0.30**
Min						**1.49**	**5.36**	**0.11**
Max						**19.76**	**14,200.00**	**0.49**

[1]As at 31/12/2023

Note: This table presents the profiles of ETFs, which include their symbol, name, benchmark, provider, replication method, inception date, age as at 31/12/2023, net assets as at 31/12/2023, and expense ratio.

Table 2 includes the descriptive statistics of ETFs' and underlying benchmarks' daily returns. Return has been calculated in raw terms by dividing the difference between the close trade price of each ETF on day t and day t-1 by the close trade price on day t-1. The return of benchmarks has been calculated in the same way with daily close prices. The descriptive statistics are presented over the entire trading history of each ETF. The average daily return of ETFs and benchmarks is 2.4 and 2.8 bps, respectively. The median return of ETFs is higher than that of benchmarks (3.3 bps vs 1.3 bps, respectively). The average risk of ETFs is equal to 1.497, being slightly lower than the average risk of benchmarks. The average extreme returns of ETFs (and benchmarks) range from -9.39% (-11.25%) to 8.68% (10.63%). At the historical cumulative level, the average return of ETFs is 16.1%. The corresponding average return of benchmarks is 17.9%.

Table 2: Descriptive Statistics of Returns

Panel A: ETFs									
Symbol	**Average %**	**Median %**	**StDev %**	**Min %**	**Max %**	**Tot.Ret. %**	**Skew**	**Kurt**	**Obs**
RMAU	0.032	0.061	0.996	-4.937	5.492	29.569	-0.197	3.518	976
XGLD	0.014	0.019	0.972	-8.350	5.865	35.558	-0.391	4.964	3,212
XGDU	0.024	0.035	0.922	-4.927	4.110	20.370	-0.370	2.847	930
XPPT	0.043	0.077	1.920	-8.217	6.535	25.433	-0.186	0.644	934
XSLR	0.065	0.029	1.866	-9.749	7.375	54.999	0.005	2.460	925
SGLD	0.014	0.020	0.973	-8.245	5.874	35.999	-0.378	5.127	3,212
SPPT	0.030	0.143	2.003	-12.305	12.650	10.706	-0.328	3.863	1,037
SSLV	-0.001	0.014	1.820	-11.560	9.629	-43.417	-0.425	5.028	3,208
IGLN	0.015	0.021	0.974	-8.481	5.853	36.816	-0.430	5.385	3,212
IGLG	0.035	0.016	0.852	-2.838	2.847	12.300	0.239	1.263	374
IPDM	0.032	0.000	2.148	-18.802	19.598	34.677	-0.009	8.158	3,212
IPLT	0.030	0.111	2.003	-11.761	12.366	10.829	-0.307	3.515	1,037
ISLN	-0.001	0.022	1.821	-11.597	9.470	-43.442	-0.432	4.838	3,212
ICOM	0.021	0.034	0.953	-4.737	4.569	31.750	-0.405	3.263	1,629
ROLL	0.029	0.053	0.965	-5.136	4.636	37.393	-0.413	3.367	1,322
EXXY	-0.003	0.000	0.990	-6.570	7.158	-26.867	-0.175	3.306	4,141
IGLD	0.032	-0.023	0.843	-2.430	3.153	11.191	0.456	1.264	375
GBS	0.014	0.017	0.979	-7.978	5.835	33.260	-0.323	4.966	3,212
PHAG	-0.002	0.032	1.816	-11.158	9.676	-44.715	-0.427	4.801	3,212
PHAU	0.014	0.017	0.973	-8.176	5.945	33.533	-0.347	5.067	3,212
PHPD	0.032	0.055	2.148	-18.004	19.129	32.491	0.048	8.236	3,212
PHPT	0.029	0.127	2.021	-11.850	15.093	9.650	-0.169	5.202	1,037
WGLD	0.024	0.032	0.884	-4.584	4.228	16.831	-0.095	2.391	775
COTN	0.020	0.000	1.803	-10.781	10.744	17.804	0.231	3.466	4,040
COPA	0.018	0.000	1.959	-10.863	12.328	-5.601	0.089	2.494	4,040
COFF	0.055	0.000	2.033	-8.470	9.011	52.498	0.213	0.945	1,239
SUGA	0.013	0.000	1.869	-11.646	8.524	-16.191	0.002	1.596	4,040
BRND	0.047	0.000	2.425	-19.005	15.299	48.600	-0.068	7.867	2,209
Average	**0.024**	**0.033**	**1.497**	**-9.398**	**8.678**	**16.144**	**-0.164**	**3.923**	**2,256**
Min	**-0.003**	**-0.023**	**0.843**	**-19.005**	**2.847**	**-44.715**	**-0.432**	**0.644**	**374**
Max	**0.065**	**0.143**	**2.425**	**-2.430**	**19.598**	**54.999**	**0.456**	**8.236**	**4,141**

Panel B: Benchmarks									
Symbol	**Average %**	**Median %**	**StDev %**	**Min %**	**Max %**	**Tot.Ret. %**	**Skew**	**Kurt**	**Obs**
RMAU	0.032	0.026	1.005	-5.128	5.267	29.546	-0.239	3.304	976
XGLD	0.015	0.014	0.987	-9.150	5.267	40.347	-0.385	5.640	3,212
XGDU	0.024	0.026	0.933	-5.128	3.552	19.765	-0.277	2.287	930
XPPT	0.050	0.082	2.573	-9.834	9.847	17.350	0.007	2.701	934
XSLR	0.068	0.000	1.946	-8.734	10.748	56.874	0.392	3.277	925
SGLD	0.015	0.014	0.987	-9.150	5.267	40.347	-0.385	5.640	3,212
SPPT	0.035	0.000	2.604	-13.427	10.939	0.946	-0.056	3.336	1,037
SSLV	0.002	0.000	1.927	-17.787	18.963	-41.751	-0.268	12.449	3,208
IGLN	0.015	0.014	0.987	-9.150	5.267	40.347	-0.385	5.640	3,212
IGLG	0.040	-0.002	0.857	-3.319	3.156	16.011	0.311	1.562	374
IPDM	0.033	0.000	2.129	-14.510	18.485	40.226	-0.255	6.984	3,212
IPLT	0.035	0.000	2.604	-13.427	10.939	0.946	-0.056	3.336	1,037
ISLN	0.002	0.000	1.927	-17.787	18.963	-40.850	-0.268	12.436	3,212
ICOM	0.023	0.064	0.981	-6.059	7.465	34.117	-0.325	5.448	1,629
ROLL	0.032	0.062	0.947	-5.523	3.564	40.383	-0.544	2.993	1,322
EXXY	0.004	0.000	1.204	-15.903	17.538	-13.400	0.634	53.006	4,141
IGLD	0.035	-0.006	0.857	-3.307	3.154	12.560	0.316	1.552	375
GBS	0.015	0.014	0.987	-9.150	5.267	40.347	-0.385	5.640	3,212
PHAG	0.002	0.000	1.927	-17.787	18.963	-40.850	-0.268	12.436	3,212
PHAU	0.015	0.014	0.987	-9.150	5.267	40.347	-0.385	5.640	3,212
PHPD	0.033	0.000	2.129	-14.510	18.485	40.226	-0.255	6.984	3,212
PHPT	0.035	0.000	2.604	-13.427	10.939	0.946	-0.056	3.336	1,037
WGLD	0.023	0.008	0.889	-4.410	3.208	16.232	-0.046	1.698	775
COTN	0.022	0.000	1.875	-23.885	9.064	17.596	-0.689	9.051	4,040
COPA	0.020	0.000	1.671	-10.744	12.500	-4.213	0.057	3.998	4,040
COFF	0.071	0.000	2.235	-8.626	10.028	57.809	0.317	1.119	1,239
SUGA	0.038	0.000	2.074	-11.632	13.953	-13.799	0.097	3.110	4,040
BRND	0.048	0.045	2.613	-24.404	31.547	51.921	0.320	17.492	2,209
Average	**0.028**	**0.013**	**1.623**	**-11.252**	**10.629**	**17.869**	**-0.110**	**7.218**	**2,256**
Min	**0.002**	**-0.006**	**0.857**	**-24.404**	**3.154**	**-41.751**	**-0.689**	**1.119**	**374**
Max	**0.071**	**0.082**	**2.613**	**-3.307**	**31.547**	**57.809**	**0.634**	**53.006**	**4,141**

Note: This table presents the descriptive statistics of ETFs and benchmarks' returns, which include average and median daily returns, standard deviation of returns, minimum and maximum values, and the skewness and kurtosis estimates. Total (cumulative) returns over the entire trading history of each ETF are presented too.

2.2 Research Methods

First, we evaluate the performance of commodity ETFs via the following time series regression model:

$$R_{cp,i} = \alpha_0 + \beta_1 R_{b,i} + \upsilon \qquad (1)$$

where $R_{cp,i}$ is the daily return of the commodity ETF i and $R_{b,}$ is the daily return of the underlying commodity or commodity index i. If the examined ETFs are fully aligned with the underlying assets, alphas will be statistically insignificant, while beta will be close to unity.

After running model (1) for each ETF in the sample, we compute tracking errors in four ways found in Frino and Gallagher (2001). The first method (TE_1) regards the average daily difference in returns between ETFs and benchmarks. The second method (TE_2) concerns the total (cumulative) tracking error over the entire trading history of each ETF. The third method (TE_3) regards the standard deviation in return differences between ETFs and benchmarks. The fourth method (TE_4) concerns the standard errors of the performance regression model (1)[2].

[2] According to Frino and Gallagher (2001), tracking errors obtained from the third and the fourth method will approximate each other provided that betas estimated by model (1) will be close to unity.

In the next step, we assess the impact on tracking error by the possible departure of ETFs from a full replication policy by running the following cross-sectional regression model:

$$TE = \lambda_0 + \lambda_1 NFR + \lambda_2 ReplMet + \lambda_3 Age + \lambda_4 ExpRatio + \lambda_5 Assets + u \quad (2)$$

where TE is the tracking error estimated via methods 1 to 4, NFR (non-full replication) is the difference between model's (1) betas from unity, ReplMet refers to replication method, which is a dummy variable taking zero value when the ETF applies physical replication and 1 when the ETF pursues synthetic replication, Age is the natural logarithm of ETFs' age as at 31/12/2023, ExpRatio is the expense ratio of ETFs, and Assets regard the natural logarithm of ETFs' assets as at 31/12/2023.

In this model, we assume that the larger the gap between beta and unity, the highest the tracking error of ETFs, either positive or negative. In addition, based on findings in the literature (e.g., Fasas, 2014, and Rompotis, 2016), the physically backed ETFs are expected to have lower tracking error than synthetic ETFs. Thus, the ReplMet (replication method) coefficient is expected to be positive, as the constant of the model captures the tracking error of the physically backed ETFs and λ2 indicates the difference in tracking errors between synthetic and physical ETFs. Furthermore, as age can reflect the accumulated experience and skill of an ETF's manager, the relevant coefficient in model (2) is expected to be negative, indicating that the oldest the ETF, the lowest its tracking error. Moreover, as expenses are considered to be one of the major causes of tracking error (Frino and Gallagher, 2001, and Chu, 2011), the correlation between expense ratio and tracking error in model (2) should be positive. Finally, Chu (2011) and Drenovak et al. (2014) find that the size (assets) of a fund is negatively related the fund's tracking error indicating that big funds are more capable of tracking their benchmark than small funds. Thus, the coefficient of assets in model (2) is expected to be negative.

Along with the assessment of the impact on tracking error by the factors included in model (2), we examine if (and how) the size of the tracking error depends on the ETFs' managing company. We do so by applying the following cross-sectional regression model:

$$TE = \lambda_0 + \lambda_1 DWS + \lambda_2 Invesco + \lambda_3 iShares + \lambda_4 WisdomTree + u \quad (3)$$

where TE is defined as above. DWS is a dummy variable with value of 1 when the ETF is provided by DWS and zero otherwise. Invesco is a dummy variable with value of 1 when the ETF is provided by Invesco and zero otherwise. iShares is a dummy variable taking value 1 when the ETF is provided by BlackRock and zero otherwise. Finally, Wisdom Tree is a dummy variable with value 1 when the ETF is provided by Wisdom Tree and zero otherwise. The constant of the model captures the tracking error of the one ETF managed by HANETF. Significant differences in tracking errors among the managing firms are to be verified by statistically significant coefficients of the dummy variables.
In the last step, we analyse further the tracking error of the examined commodity ETFs by summing for each ETF the number of days with nil tracking error, negative tracking error and positive tracking error, respectively.

3. Results

The results of model (1) on the performance of commodity ETFs are presented in Table 3. The average alpha of the sample is actually nil. In addition, with no exceptions, alphas are not statistically significant. This finding is not surprising as the examined ETFs do not seek to beat their underlying commodities and indexes.

Table 3: Descriptive Statistics of Returns

Symbol	alpha	t-stat[1]	beta	t-stat[2]	R-2	Obs	NFR
RMAU	0.005	0.283	0.851[a]	-9.129	0.737	976	0.149
XGLD	0.001	0.134	0.844[a]	-17.367	0.734	3,212	0.156
XGDU	0.004	0.264	0.844[a]	-9.258	0.729	930	0.156
XPPT	0.024	0.444	0.666[a]	-15.680	0.413	934	0.334
XSLR	0.028	0.553	0.546[a]	-17.491	0.324	925	0.454
SGLD	0.001	0.134	0.853[a]	-16.884	0.748	3,212	0.147
SPPT	0.016	0.302	0.597[a]	-19.664	0.467	1,037	0.403
SSLV	-0.002	-0.081	0.552[a]	-33.093	0.342	3,208	0.448
IGLN	0.001	0.161	0.849[a]	-16.995	0.740	3,212	0.151
IGLG	-0.001	-0.027	0.814[a]	-6.278	0.671	374	0.186
IPDM	0.008	0.304	0.735[a]	-21.731	0.531	3,212	0.265
IPLT	0.016	0.305	0.595a	-19.741	0.464	1,037	0.405
ISLN	-0.002	-0.090	0.554[a]	-33.025	0.343	3,212	0.446
ICOM	0.003	0.230	0.808[a]	-14.357	0.692	1,629	0.192
ROLL	0.003	0.208	0.851[a]	-9.648	0.697	1,322	0.149
EXXY	-0.005	-0.558	0.659[a]	-44.589	0.643	4,141	0.341
IGLD	-0.002	-0.164	0.953[a]	-3.745	0.940	375	0.047
GBS	0.001	0.066	0.852[a]	-16.496	0.739	3,212	0.148
PHAG	-0.003	-0.122	0.554[a]	-33.155	0.346	3,212	0.446
PHAU	0.001	0.072	0.850[a]	-16.988	0.743	3,212	0.150
PHPD	0.007	0.279	0.740a	-21.484	0.538	3,212	0.260
PHPT	0.015	0.286	0.600[a]	-19.362	0.465	1,037	0.400
WGLD	0.004	0.250	0.850[a]	-8.075	0.730	775	0.150
COTN	0.006	0.290	0.653[a]	-31.244	0.462	4,040	0.347
COPA	0.012	0.389	0.651[a]	-19.615	0.698	4,040	0.349
COFF	0.000	-0.007	0.769[a]	-16.737	0.716	1,239	0.231
SUGA	-0.012	-0.645	0.680[a]	-34.398	0.569	4,040	0.320
BRND	0.037	0.738	0.622[a]	-19.700	0.455	2,209	0.378
Average	**0.006**	**0.143**	**0.728**	**-19.497**	**0.596**	**2,256**	**0.272**
Min	**-0.012**	**-0.645**	**0.546**	**-44.589**	**0.324**	**374**	**0.047**
Max	**0.037**	**0.738**	**0.953**	**-3.745**	**0.940**	**4,141**	**0.454**

[1] t-stat for alphas being statistically different from zero; [2] t-stat for betas being statistically different from unity
[a] Statistically significant at 1%
NFR= Non Full Replication as evidenced by the differences between ETFs' betas and unity.
Note: This table presents the results of a single factor time series regression model in which the daily return of each ETF is regressed on the corresponding return of its benchmark.

The average beta is 0.73 indicating that the sample's commodity ETFs are quite aligned to their tracking assets. However, by focusing on the single beta estimates, we see that all beta estimates are statistically different from unity. Overall, betas indicate that the examined UK-listed commodity ETFs are not fully aligned with their underlying benchmarks. This departure from the full alignment (amounting to 0.27 on average as shown in Table 3) may be indicative of significant tracking errors.

Indeed, as we see in Table 4, the examined commodity ETFs fail to fully replicate the performance of their benchmarks. At the daily level, the average tracking error of the sample is slightly negative at -0.5 bps. Twenty four out of the twenty eight ETFs present negative tracking error. This negative tracking error indicates that the corresponding ETFs underperform their benchmarks.

Table 4: Measures of Tracking Error

Ticker	TE_1 (Average)	TE_2 (Total)	TE_3 (StDev)	TE_4 (SE)	Min	Max	Obs
RMAU	0.000	0.022	0.532	0.511	-2.682	3.189	976
XGLD	-0.001	-4.789	0.525	0.502	-3.395	3.875	3,212
XGDU	0.000	0.605	0.502	0.480	-2.312	2.237	930
XPPT	-0.008	8.083	1.335	1.073	-10.015	8.698	934
XSLR	-0.003	-1.876	1.770	1.534	-7.162	7.984	925
SGLD	-0.001	-4.348	0.509	0.488	-3.397	3.572	3,212
SPPT	-0.005	9.761	1.324	1.172	-9.983	12.144	1,037
SSLV	-0.003	-1.667	1.709	1.476	-11.839	9.869	3,208
IGLN	-0.001	-3.531	0.519	0.497	-3.394	3.659	3,212
IGLG	-0.009	-3.711	0.514	0.489	-2.361	2.236	374
IPDM	-0.001	-5.549	1.576	1.472	-11.997	14.994	3,212
IPLT	-0.005	9.884	1.332	1.072	-10.092	12.200	1,037
ISLN	-0.003	-2.591	1.708	1.476	-11.904	10.063	3,212
ICOM	-0.001	-2.367	0.562	0.529	-6.865	5.848	1,629
ROLL	-0.001	-2.990	0.549	0.531	-3.965	4.121	1,322
EXXY	-0.006	-13.467	0.720	0.592	-17.382	14.746	4,141
IGLD	-0.003	-1.369	0.210	0.206	-2.109	1.408	375
GBS	-0.002	-7.087	0.521	0.500	-3.279	3.515	3,212
PHAG	-0.004	-3.865	1.702	1.469	-12.597	10.270	3,212
PHAU	-0.002	-6.814	0.515	0.493	-3.345	3.832	3,212
PHPD	-0.001	-7.734	1.560	1.459	-12.173	14.526	3,212
PHPT	-0.006	8.704	1.133	0.933	-9.975	11.873	1,037
WGLD	0.001	0.599	0.478	0.459	-2.332	2.234	775
COTN	-0.002	0.208	1.473	1.322	-7.900	9.431	4,040
COPA	-0.002	-1.388	0.914	0.790	-10.733	11.023	4,040
COFF	-0.017	-5.312	1.200	1.084	-6.826	11.316	1,239
SUGA	-0.025	-2.392	1.395	1.227	-12.071	10.485	4,040
BRND	0.000	-3.321	1.219	1.057	-11.753	10.651	2,209
Average	**-0.005**	**-3.199**	**0.991**	**0.883**	**-8.451**	**8.459**	**2,256**
Min	**-0.025**	**-13.467**	**0.210**	**0.206**	**-17.382**	**1.408**	**374**
Max	**0.001**	**8.704**	**1.708**	**1.476**	**-2.109**	**14.746**	**4,141**

Note: This table presents the tracking error of ETFs. Tracking error is calculated in four alternative ways, that is, i) average daily return difference between ETFs and benchmarks, ii) total (cumulative) tracking errors over the entire trading history of each ETF, iii) standard deviation in daily return differences between ETFs and benchmarks, and iv) sum of standard errors (SE) deriving from the single factor regression model where the daily return of each ETF is regressed on the corresponding return of its benchmark, Extreme daily tracking errors are reported too.

The underperformance of ETFs is more evident when cumulative tracking errors are taken into consideration. The respective average term is 3.2% (or 320 bps). Maximum underperformance is -13.5%, while maximum outperformance is 8.7%. These extreme tracking errors are shown by the iShares Diversified Commodity Swap UCITS ETF (DE) and the WisdomTree Physical Platinum, i.e., two of the ETFs that significantly depart from the full replication, as inferred by their betas which significantly decline from unity.

The next two methods used to calculate tracking error also indicate that the return gap between commodity ETFs and their benchmarks is significant. The average TE_3 of the sample is equal to 99 ps. The average TE4 equals 88 bps. To some extent, the difference of 11 bps between the average TE3 and TE_4 tracking error figures must be the result of beta estimates in Table (3) being lower than unity by an average of 27 bps. Other factors can explain tracking error too.

In fact, as reported in Table 5, the coefficients of NFR are positive and statistically significant for TE2, TE3 and TE4, verifying our expectations about a positive correlation between tracking error and the departure from the full replication strategy. The results on the dummy concerning the replication

method are also in agreement with our assumption about physical ETFs being more efficient in replicating their benchmarks compared to their synthetic peers. In particular, the relevant coefficients for TE3 and TE4 are significantly positive indicating that the synthetic ETFs have higher tracking error than the physically backed ETFs.[3]

Table 5: Tracking Error Per Factors Regression Results

	Dep. Var.: TE1		Dep. Var.: TE2		Dep. Var.: TE3		Dep. Var.: TE4	
	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat
Constant	0.01	1.53	-0.51	-0.04	1.35[a]	3.18	1.35[a]	3.25
NFR	0.00	0.22	19.67[c]	1.89	2.80[a]	6.98	2.13[a]	5.42
Repl. Method	0.00	-0.53	-2.37	-1.07	0.22[b]	2.14	0.19[c]	1.94
Age	0.00	0.22	-1.87	-1.06	0.16[b]	2.20	0.16[b]	2.33
Expense Ratio	0.01[b]	2.07	-3.38	-0.32	0.01	0.02	-0.06	-0.16
Assets	0.00	-1.44	-0.04	-0.06	-0.07[a]	-3.51	-0.07[a]	-3.40
R-2	0.34		0.26		0.85		0.81	
Obs	28		28		28		28	

[a] Statistically significant at 1%; [b] Statistically significant at 5%; [c] Statistically significant at 10%

Note 1: This table presents the results of a cross-sectional regression model in which the tracking errors of ETFs is regressed on their non-full replication policy (NFR) as evidenced by the differences between their betas (in Table 3) and unity, replication method, that is, a dummy variable taking zero value when the ETF applies physical replication and one when the ETF pursues synthetic replication, age as at 31/12/2023, expense ratio, and assets as at 31/12/2023.

Note 2: The absolute value of TE_1 and TE_2 is used in this model.

Going further, the coefficients of age in Table 5 are significantly positive for TE3 and TE4 and insignificant for TE1 and TE2. The significantly positive estimates for age contradict our assumption about the positive impact on the tracking ability of an ETF exerted by the accumulated experience of the ETF's manager as the latter may be reflected by the age of ETFs. Moreover, our assumption about the positive correlation between tracking error and expense ratios is verified only for TE1. Finally, our expectation about the negative relationship between tracking error and the magnitude of ETFs' assets is verified. In particular, the estimates of the assets factor are significantly negative for TE3 and TE4.

Based on the regression results, we can conclude that the five determinative factors included in model (2) are quite capable of explaining the tracking error of the UK-listed commodity ETFs. This ability is verified by the relatively high R-squared values, especially for TE3 and TE4. However, this is not the case when assessing the impact on tracking error by the individual providers of commodity ETFs included in model (3). As shown in Table (6), all the relevant estimates are statistically insignificant indicating that there are no statistically and economically differences in the tracking efficiency among the five managing companies considered in our analysis.

[3] The average TE3 (TE4) of physical ETFs amounts to 0.999% (0.888%). The corresponding figures for synthetic ETFs are 1.004% and 0.892%.

Table 6: Tracking Error Per Provider Regression Results

	Dep. Var.: TE1		Dep. Var.: TE2		Dep. Var.: TE3		Dep. Var.: TE4	
	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat
Constant	0.00	0.01	0.02	0.00	0.53	1.03	0.51	1.17
DWS	0.00	0.47	0.48	0.08	0.50	0.86	0.39	0.79
Invesco	0.00	0.46	1.23	0.19	0.65	1.08	0.53	1.06
iShares	0.00	0.58	-2.88	-0.49	0.32	0.59	0.25	0.54
Wisdom Tree	0.01	0.93	-2.60	-0.45	0.57	1.05	0.47	1.03
R-2	0.07		0.09		0.09		0.09	
Obs	28		28		28		28	

Note 1: This table presents the results of a cross-sectional regression model in which the constant expresses the ETFs managed by HANETF and four dummy variables for ETFs managed by DWS, Invesco, iShares (BlackRock), and Wisdom Tree, respectively.
Note 2: The absolute value of TE_1 and TE_2 is used in this model.

The decomposition of daily tracking errors is presented in Table 7. More specifically, the table shows that, on average, ETFs achieve zero tracking errors just in 0.35% of total trading days. Positive tracking errors are computed in about 47% of trading days. The average positive daily tracking error (outperformance) amounts to 84 bps. On the other hand, negative tracking errors are realised in about 52% of trading days. The average negative daily tracking error is equal to -84 bps. In sum, the ETFs under study underperform their benchmarks slightly more frequently than they outperform them.

Table 7: Analysis of Daily Tracking Error

Ticker	Nil TE	% Nil TE	Posit. TE	% Pos. TE	Av. Posiv. TE	Neg. TE	% Neg. TE	Av. Neg. TE	Obs
RMAU	0	0.00%	486	49.80%	0.396	490	50.20%	-0.392	976
XGLD	1	0.03%	1,593	49.60%	0.375	1,618	50.37%	-0.371	3,212
XGDU	0	0.00%	463	49.78%	0.385	467	50.22%	-0.381	930
XPPT	0	0.00%	470	50.32%	1.570	464	49.68%	-1.610	934
XSLR	0	0.00%	470	50.81%	1.359	455	49.19%	-1.409	925
SGLD	1	0.03%	1,602	49.88%	0.367	1,609	50.09%	-0.368	3,212
SPPT	0	0.00%	524	50.53%	1.554	513	49.47%	-1.597	1,037
SSLV	0	0.00%	1,590	49.56%	1.262	1,618	50.44%	-1.246	3,208
IGLN	1	0.03%	1,591	49.53%	0.375	1,620	50.44%	-0.370	3,212
IGLG	0	0.00%	181	48.40%	0.380	193	51.60%	-0.373	374
IPDM	8	0.25%	1,611	50.16%	1.127	1,593	49.60%	-1.141	3,212
IPLT	1	0.10%	523	50.43%	1.563	513	49.47%	-1.603	1,037
ISLN	1	0.03%	1,592	49.56%	1.260	1,619	50.40%	-1.246	3,212
ICOM	4	0.25%	818	50.21%	0.386	807	49.54%	-0.393	1,629
ROLL	1	0.08%	667	50.45%	0.381	654	49.47%	-0.391	1,322
EXXY	37	0.89%	610	14.73%	0.419	3,494	84.38%	-0.080	4,141
IGLD	0	0.00%	79	21.07%	0.078	296	78.93%	-0.025	375
GBS	1	0.03%	1,583	49.28%	0.377	1,628	50.68%	-0.370	3,212
PHAG	4	0.12%	1,579	49.16%	1.263	1,629	50.72%	-1.233	3,212
PHAU	2	0.06%	1,593	49.60%	0.371	1,617	50.34%	-0.369	3,212
PHPD	5	0.16%	1,611	50.16%	1.113	1,596	49.69%	-1.126	3,212
PHPT	0	0.00%	523	50.43%	1.556	514	49.57%	-1.594	1,037
WGLD	0	0.00%	378	48.77%	0.374	397	51.23%	-0.356	775
COTN	14	0.35%	2,003	49.58%	0.974	2,023	50.07%	-0.967	4,040
COPA	262	6.49%	1,900	47.03%	1.611	1,878	46.49%	-1.634	4,040
COFF	2	0.16%	616	49.72%	0.830	621	50.12%	-0.856	1,239
SUGA	10	0.25%	2,036	50.40%	0.885	1,994	49.36%	-0.953	4,040
BRND	8	0.36%	1,095	49.57%	1.040	1,106	50.07%	-1.010	2,209
Average	**13**	**0.35%**	**1,064**	**47.45%**	**0.844**	**1,180**	**52.21%**	**-0.838**	**2,256**
Min	**0**	**0.00%**	**79**	**14.73%**	**0.078**	**193**	**46.49%**	**-1.634**	**374**
Max	**262**	**6.49%**	**2,036**	**50.81%**	**1.611**	**3,494**	**84.38%**	**-0.025**	**4,141**

Note: This table presents an analysis of ETFs' daily tracking error. This analysis considers the number of days where ETFs present zero tracking error, the number of days where tracking error is positive, and the number of days where tracking error is negative.

A last comment that should be made with respect to the tracking efficiency of ETFs, is that, despite the presence of ETFs for about three decades now (given that the US-listed SPDRs tracking the S&P 500 Index was the first ETF to enter the stock markets worldwide in 1993), tracking inefficiencies are still there, as they used to be during the first years of ETFs' existence. These inefficiencies must relate to inherent frictions attached to the passively managed ETFs which try to replicate the return of benchmarks which are not affected by expenses, age, assets and other factors that affect the replication efforts of ETFs. These frictions have been accentuated by several studies in literature and are confirmed by the current study too.

4. Conclusion

The performance and tracking efficiency of twenty eight commodity ETFs that are traded on the London Stock Exchange are examined in this study. The analysis shows that the average daily and cumulative return of these ETFs has been positive during their entire trading history. However, the return of ETFs has been inferior to the return of their underlying commodities and indexes by 320 bps, indicating a significant tracking inefficiency. Tracking inefficiency is verified by all the methods used to compute the tracking error of the examined commodity ETFs.

One key factor that can provoke tracking inefficiency relates to the inability of ETFs to be fully aligned with their underlying assets. Non-full alignment might also be a choice made by the examined commodity ETFs. In any case, the departure from the full replication is inferred by the fact that the beta estimates obtained from the performance regression model differ statistically from unity in eight out of nine cases. By relevant regression analysis, it is verified that the non-full alignment to underlying benchmarks is positively related to the tracking error of ETFs, which, by the way, is negative on about 52% of days over the entire trading history of commodity ETFs in the UK. Other factors that can induce tracking error include the replication method applied by ETFs, their age, assets, and, to a less degree, their expense ratio.

References

Chu, P.K.K., (2011). Study on the Tracking Errors and their Determinants: Evidence from Hong Kong Exchange Traded Funds. *Applied Financial Economics* 21(5), pp. 309-315.

Drenovak, M., Urošević, B., and Jelic, R. (2014). European Bond ETFs: Tracking Errors and the Sovereign Debt Crisis. *European Financial Management* 20(5), pp. 958-994.

Fassas, A.P. (2014). Tracking Ability of ETFs: Physical versus Synthetic Replication. *Journal of Index Investing* 5(2), pp. 9-20.

Frino, A. & Gallagher D.R. (2001). Tracking S&P 500 Index Funds. *Journal of Portfolio Management* 28(1), pp. 44-55.

Guedj, I., Li, G., & McCann, C. (2011). Futures-Based Commodities ETFs. *Journal of Index Investing* 2(1), pp. 14-24.

Guo, K., & Leung, T. (2015). Understanding the Tracking Errors of Commodity Leveraged ETFs. *Commodities, Energy and Environmental Finance*, Springer, pp. 39-63.

Mateus, C., & Rahmani, Y. (2015). Physical versus Synthetic Exchange Traded Funds. Which one Replicates Better?. *Journal of Mathematical Finance* 7, pp. 975-989.

Maurer, F., & Williams, O. (2015). Physically Versus Synthetically Replicated Trackers: Is There A Difference In Terms Of Risk?. *Journal of Applied Business Research,* 31(1), pp. 131-146.

Merz, T. (2015). The Tracking Risk of Exchange-Traded Funds Revisited: A Multivariate Regression Approach. Working Paper, Zurich University of Applied Sciences, School of Management and Law.

Murphy, R., & Wright, C. (2010). An Empirical Investigation of the Performance of Commodity-Based Leveraged ETFs. *Journal of Index Investing* 1(3), pp. 14-23.

Neff, T., & Isengildina-Massa, O. (2018). How Well Do Commodity ETFs Track Underlying Assets?. Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Minneapolis, MN.

Perera, D., Białkowski, J., & Bohl, M.T. (2022). Is the Tracking Error Time-Varying? Evidence from Agricultural ETCs. *Research in International Business and Finance* 63, pp. 1-21.

Rompotis, G.G. (2016). Physical versus Futures-Based Replication: The Case of Commodity ETFs. *Journal of Index Investing* 7(2), pp. 16-37.

Sousa, J.M.L. (2014). Tracking Ability of Metal Exchange Traded Funds (ETFs). Working Paper, ISCTE Business School.

Stewart, S.L., Isengildina-Massa, O., Hassman, C., & de Leon, M. (2023). ETP Tracking of U.S. Agricultural and Energy Markets. *Journal of Commodity Markets* 31, pp. 1-16.