Wei-Fong Pan (China), Ting Li (China)

The measurement of tracking errors of commodity ETFs in China

Abstract

This paper presents the first study on the measurement and determinants of tracking errors using the daily figures for gold exchange-traded funds (ETFs) in China. This study employs three methods to measure tracking errors – one that involves calculating the absolute error measure, one that involves calculating the differences between the standard deviation of the benchmark index and that of the ETF, and a regression analysis of empirical returns. In general, the results suggest that the tracking errors of these ETFs in China are lower than those of equity-based ETFs in Hong Kong, the United States, and Australia. We also observe that distinct ETFs have different determinants. Our results provide valuable insight for both institutional and retail investors, as well as opportunities for them to be exposed to a wide range of commodity ETFs in China.

Introduction

The development of exchange-traded funds (ETFs) provides opportunities for both institutional and retail investors to be exposed to a wide range of asset classes.

A bulk of the existing studies focus on the tracking errors of equity-based ETFs using distinct approaches. Using S&P 500 Index data, Frino et al. (2004) examined the exogenous determinants of tracking errors and observed that such errors are significantly influenced by index revisions, share issuances, spin-offs, share repurchases, index replication strategy and fund size. They also found a seasonal pattern in tracking errors, consistent with the finding of Frino and Gallagher (2001). Chu (2011) studied the magnitude and determinants of ETF tracking errors using daily data in the Hong Kong stock market and found that the tracking error in Hong Kong is higher than those in the United States and Australia. Avellaneda and Zhang (2010) studied the price behavior equity-leveraged ETFs in different sectors and found minimal one-day tracking errors among the most liquid equity ETFs.

Commodities are unique in part because physical assets cannot be stored easily owing to the extra costs for warehousing. Thus, futures-based commodity ETFs may fail to track their reference indices perfectly. The commodity is also counter-cyclical with stocks and bonds; studies observed that it is significantly negatively correlated with both bonds and equities, implying that an appropriate allocation to commodities enhances portfolio performance (Jensen et al., 2002; Fuertes et al., 2010). Gold is often viewed by investors as a

hedge against market turmoil. A typical example of this tendency is when the price of gold was pushed to an all-time high of US$1,900 in August 2011 owing to the global financial crisis and European sovereign debt crisis at the time[1]. Although commodities, especially gold, are important both for risk hedging and for asset management, studies of tracking errors in commodity ETFs remain few[2].

The contribution of this study is to measure the determinants and magnitude of the tracking error for commodity ETFs in China, from 05 January 2015 to 29 February 2016. To the best our knowledge, this study is the first to investigate all four existing gold ETFs in China. Existing studies paid more attention to equity-based ETFs in either the United States or European countries rather than in emerging countries. Following Pope and Yadav (1994) and Shin and Soydemir (2010), this study employed three different approaches to estimate tracking errors in order to obtain robust results.

The rest of the paper is organized as follows. Section 2 presents the data sources and provides an overview of the development of commodity ETFs in China. Section 3 describes the empirical approaches used to estimate the tracking errors as well as its determinants. Section 4 discusses the empirical findings, and Section 5 provides the conclusions and some directions for future research.

1. The development of commodity ETFs in China

The development of gold ETFs enables investors to allocate some of their assets to gold without directly buying physical gold. Gold ETFs in China first emerged on 24 June 2013, developed by GuoTai

---

Wei-Fong Pan, Investment Analyst, Sales and Trading Department, PingAn Futures Co., Ltd, China.
Ting Li, Portfolio Manager, Sino Life Asset Management Co., China.

[1] Białkowski et al. (2015) investigated the gold price during these crises.
[2] Some of these few existing studies are those by Murphy and Wright (2010), Guedj et al. (2011) and Leung and Ward (2015).

Fund Management Company, and the country has since become the largest gold consumer in the world. The Shanghai Gold Exchange facilitates spot gold exchange.

Table 1 shows that the trading volume of spot gold in China has significantly increased along with the trading amount, suggesting that investors have become more focused on gold investments, which, in turn, makes this study important and timely.

The commodity ETFs used in this study are HuaAn Gold ETF, GuoTai Gold ETF, Bosera Gold ETF, and E Fund Gold ETF[1]. The ETF prices were collected from the Wind Database, created by Wind Information Co., Ltd., a financial data provider in China. Since the commodity ETFs in China emerged later than those in developed countries, all four commodity ETFs track the gold spot price at the Shanghai Gold Exchange, which is also the source of the gold spot price in this study. All of the data reflect daily observations for each trading day from 05 January 2015 to 29 February 2016.

Figure 1 shows the performance of the existing gold ETFs in China. All four ETFs showa similar trend, with very small variations, and have a net asset value (NAV) between 2.00 and 2.65. However, even such small variations would have a large impact on the ETF returns.

2. Three methods for tracking error estimation

This section reviews the possible sources of tracking errors and the methods for analyzing such errors. The tracking error, ceteris paribus, is zero if the index fund perfectly aligns with the benchmark index. However, in practice, an ETF's performance in tracking the index is affected by a few factors, such as management fees and administrative/operating expenses, different compositions of the index fund and the index, and trading costs (Frino and Gallagher, 2001; Drenovak et al., 2014). Thus, the tracking error is non-zero in practice, as was observed by many empirical studies (see for example, Murphy and Wright, 2010).

Several articles explored important issues in tracking error measurement. Roll (1992) provided a criterion for analyzing ETF performance. The approaches for tracking error estimation were well documented in the academic literature (e.g. Pope and Yadav, 1994; Shin and Soydemir, 2010). This study employs three methods to measure the tracking errors. One of the traditional methods

involves calculating the absolute error measure, which is defined as the average absolute value of the difference between the returns of the benchmark index and index fund. The measure can be described as follows.

$$TE_{1,i} = \frac{\left(\sum_{i=1}^{n} \left| R_{f,t} - R_{x,t} \right| \right)}{n}, \ t = 1, 2, 3 \ldots n, \qquad (1)$$

where $R_{f,t}$ represents the return of index fund f at time t, while $R_{x,t}$ is the return of its underlying gold at time t.

The second method of tracking error estimation involves calculating the standard deviation of the difference inreturns of benchmark index and that of the ETF. The variance equation can described as follows.

$$TE_{2,i} = \sqrt{ \frac{1}{n-1} \sum_{i=1}^{n} \left[\left(R_{f,t} - R_{x,t} \right) - \frac{1}{n} \sum_{i=1}^{n} \left(R_{f,t} - R_{x,t} \right) \right]^2 }, \qquad (2)$$

where t denotes the time period. $R_{f,t}$ represents the return of index fund f at time t, while $R_{x,t}$ is the return of its underlying (Gold) at time t. We can rewrite equation (2) as:

$$TE_{2,i} = \sqrt{ \frac{1}{n-1} \sum_{i=1}^{n} \left[e_{f,t} - \overline{e}_x \right]^2 }, \qquad (3)$$

where $e_{f,t} = R_{f,t} - R_{x,t}$.

The third method of tracking error estimation involves a regression analysis of empirical returns, based on the following linear model:

$$R_{f,t} = \alpha + \beta R_{x,t} + \varepsilon, \qquad (4)$$

where $\varepsilon \sim N(0, \sigma^2)$ is the error term. The tracking error is defined as the standard error of equation (4). In the case of ETFs pursuing a passive investment strategy, α is not expected to be statistically different from zero, while β is not expected to be statistically different from one. A very high R^2 is also expected.

In final step, we test the possible determinants of tracking error. Tracking-error determinants may stem from administrative/operating expenses, index fund composition variance from the index itself, trading costs, fund size, etc. (Frino and Gallagher, 2001; Drenovak et al., 2014). In this study, we employ trading volume and trading amount as cash flow indicators, and also fund size. ETF tracking- error performance may also be related to the size and timing of cash-flows and fund size. We used the following models to test the determinants of tracking errors:

$$\left| TE_{i,t} \right| = \beta_0 + \beta_1 Vol + \beta_2 Amount + \beta_3 Size + \varepsilon_{i,t}, \qquad (5)$$

---

[1] Our dataset excludes the UBS SDIC Silver LOF (listed-open fund), first, because it has been traded for less than half a year, and second, LOFs differ from ETFs in some aspects, such as in the redemption mechanism.

where $\left|TE_{i,t}\right|$ is the absolute value of tracking error measured by different definitions for fund i at time t.

3. Results

We begin with estimating the tracking error using the absolute error method (TE_1). From Figure 2, which presents the TE_1 variation of all the gold ETFs, it is clear that the highest tracking error occurs in January 2015. The reason behind this phenomenonis that the Bosera Gold ETF cannot accurately track the increasing return of spot goldin January 2015, which result in that its tracking errors are ten times higher than other three ETFs. For robustness, we used three samples according to time period – the full sample (January 2015-March 2016), the sample only for one year during the study period (March 2015-2016), and the sample only for the last six months of the study period (September 2015-March 2016). Table 2 reports the empirical results of the tracking error estimation using the three methods.

We first consider the full sample, that is, the sample for the entire study period. The daily tracking error based on the first estimation method (calculating the absolute error measure) (TE_1) ranges from 0.0024% to 0.0273% across all ETFs. The daily tracking error based on the second method (calculating the standard deviation of return differences) (TE_2) ranges from 0.0035 % to 0.05%. Meanwhile, the daily tracking error based on the third method(regression analysis of empirical returns) (TE_3) ranges from 0.0027% to 0.0499%, and the coefficient of the benchmark index, as expected, is very close to one and the R^2 and is nearly 100%. The tracking error of the gold ETFs in China is generally lower than those of equity-based index ETFs in Hong Kong (0.39%), Australia (0.0074%), and the United States (0.039%) (Chu, 2011). The measures from all three methods indicate that HuaAn Gold ETFs have highest tracking error among four ETFs and GuoTai Gold ETFs have thebest performance and the smallest tracking error.

For robustness, we also consider two other sample – one for only one year of the study period (March 2015-2016) and another for only the last six months of the study period (September 2015- March 2016). TE_1 is between 0.0032 and 0.0286 across all ETFs, TE_2 is between0.0034 and 0.0523, and TE_3 is between 0.0028 and 0.0525. For the full sample period, the order of the ETFs in terms of the magnitude of tracking error is the same when using the first two methods. However, the results of the third method (regression analysis of empirical returns) show that E Fund Gold ETFs, not GuoTai

Gold ETFs, have the lowest tracking error. A similar situation is observed for the sample for the last six months of the study period – the results of the first two methods suggest that GuoTai Gold ETFs have the smallest tracking error, but the results of the third method show that E Fund Gold ETFs have the best performance. The study's results support Pope and Yadav's (1994) idea that if β is not exactly equal to one, the order of the ETFs in terms of the magnitude of tracking error may be different. Pope and Yadav (1994) also pointed out that if the relationship between the benchmark index return and Gold ETF return is not linear, the third method may overestimate the tracking error.

Finally, we examine the determinants of tracking errors for each gold ETFs and the results are reported in Table 3. Our results show that the tracking-error determinants differ with products. Trading amount and volume generally have insignificant effects on daily tracking errors. The exception is GuoTai Gold ETFs whose trakcing-error performance displayed a negative relationship with trading volume, but positive with trading amount. The fund size is negatively correlated for HuaAn Gold ETF's tracking error, whereas it is positively correlated for that of Bosera Gold ETF. None of the three determinants had a significant impact on E Fund Gold ETF. Our results are slightly different from previous findings that show fund size as significantly negatively correlated with tracking errors of equity-based ETFs (Grinblatt and Titman, 1989; Chu, 2011). Our results conclude that various ETFs have different determinants.

Conclusion

ETFs have provided both institutional and retail investors with new opportunities to be exposed to a wide array of commodities. This study is the first to examine the measurement and determinants of tracking errors using daily data for gold ETFs in China, from January 2015 to March 2016. The study's results show that the tracking error of gold ETFs is generally lower than those of equity-based ETFs in Hong Kong, the United States and Australia. The results consistently indicate HuaAn Gold ETFs have the highest tracking error among all the four gold ETFs. The results also support Pope and Yadav (1994) finding that the tracking error calculated from regression analysis may differ from the standard deviation of return difference if the coefficient of the benchmark index is not exactly equal to one.

In regards to gold ETFs' determinants, our results conclude that various ETFs have different determinants. Trading amount and volume generally have insignificant effects on daily tracking errors.

The exception is GuoTai Gold ETFs whose trakcing-error performance displayed a negative relationship with trading volume, but positive with trading amount. The fund size is negatively correlated for HuaAn Gold ETF's tracking error, whereas it is positively correlated for that of Bosera Gold ETF. Our results are slightly different from previous findings that show fund size as significantly negatively correlated with tracking errors of equity-based ETFs (Grinblatt and Titman, 1989; Chu, 2011). A possible explanation is that commodities are different with equities and cannot be stored easily. The ETF issuers are required to trading futures with counterparties or warehousing, which is very costly (see Guedj et al., 2011). Thus, determinants of commodity ETFs are different.

These findings provide important information for investors, particularly on the measurement of commodity ETFs in China. For future research, this study's framework can be extended to investigate other types of ETFs, to investigate in the context of other countries, and to examine the other determinants of tracking errors in commodity ETFs.

References

1. Avellaneda M. and Zhang S. (2010). Path-dependence of Leveraged ETF Returns, *SIAM Journal of Financial Mathematics*, 1, pp. 586-603.
2. Białkowski, J., Bohl, M.T., Stephan, P.M. and Wisniewski, T.P. (2015). The Gold Price in Times of Crisis, *International Review of Financial Analysis*, 41, pp. 329-339.
3. Chu, P.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.
4. Frino, A. and Gallagher, D.R. (2001). Tracking S&P 500 Index Funds, *Journal of Portfolio Management*, 28(1), pp. 44-55.
5. Frino A., Gallagher, D.R., Neubert, A.S. and Oetomo, T.N. (2004). Index Design and Implications for Index Tracking: Evidence from S&P 500 Index Funds, *Journal of Portfolio Management*, pp. 88-95.
6. Fuertes, A., Miffre, J. and Rallis, G. (2010). Tactical Allocation in Commodity Futures Markets: Combining Momentum and Term Structure Signals, *Journal of Banking & Finance*, 34(10), pp. 2530-2548.
7. Grinblatt, M. and Titman, S. (1989). Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings, *Journal of Business*, 62, pp. 393-416.
8. Guedj, I., Li, G. and McCann, C. (2011). Futures-based Commodities ETFs, *The Journal of Index Investing* 2(1), pp. 14-24.
9. Jensen, G.R., Johnson, R.R. and Mercer, J.M. (2002). Tactical Asset Allocation and Commodity Futures, *The Journal of Portfolio Management*, 28(4), pp. 100-111.
10. Leung T. and Ward B. (2015). The Golden Target: Analyzing the Tracking Performance of Leveraged Gold ETFs, *Studies in Economics and Finance*, 32(3), pp. 278-297.
11. Murphy, R. and Wright, C. (2010). An Empirical Investigation of the Performance of Commodity-based Leveraged ETFs, *The Journal of Index Investing*, 1(3), pp. 14-23.
12. Pope, P. and Yadav, P. (1994). Discovering Errors in Tracking Error, *Journal of Portfolio Management*, 20, pp. 27-32.
13. Roll, R. (1992). A Mean/Variance Analysis of Tracking Error, *Journal of Portfolio Management*, 18, pp. 13-22.
14. Shin, S. and Soydemir, G. (2010). Exchange-traded Funds, Persistence in Tracking Errors and Information Dissemination, *Journal of Multinational Financial Management*, 20, pp. 214-234.
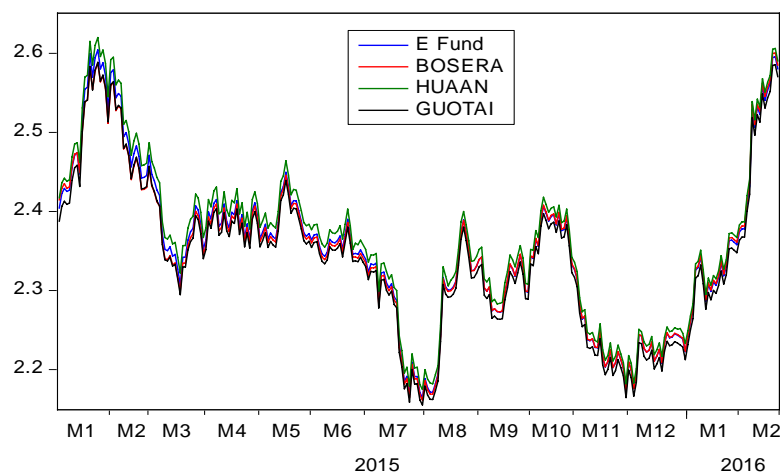
Appendix

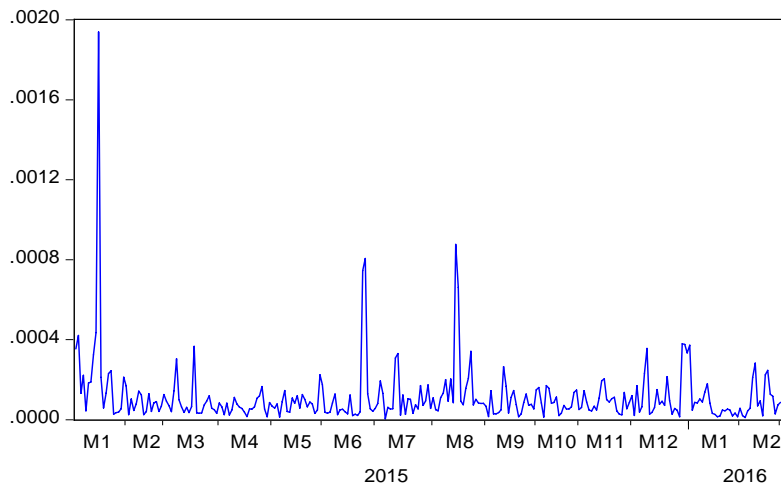**Fig. 1. Performance of the four gold ETFs in China, March 2015-2016**

Fig. 2. Average Gold ETF Tracking Errors (*TE*₁)

Table 2. Tracking errors of gold ETFs for the three samples

	Absolute error				Return differences		Regression analysis			
Full sample period										
	TE_1 (%)				TE_2 (%)		TE_3 (%)			
	Mean	SD	Min	Max	SD	Mean	ε	α	β	R^2 (%)
GuoTai	0.0024	0.0026	0.0000	0.0249	0.0035	0.0002	0.0035	0.0002	0.9996	99.9985
HuaAn	0.0273	0.0417	0.0000	0.3449	0.0500	-0.0013	0.0499	-0.0012	0.9983	99.6916
Bosera	0.0128	0.0463	0.0000	0.7089	0.0481	-0.0015	0.0466	-0.0010	0.9866	99.7241
E Fund	0.0033	0.0029	0.0000	0.0151	0.0100	-0.0009	0.0027	-0.0008	0.9962	99.9991
March 2015-2016										
GuoTai	0.0023	0.0025	0.0000	0.0249	0.0034	-0.0001	0.0034	-0.0001		0.999.9985
HuaAn	0.0286	0.0439	0.0000	0.3449	0.0523	-0.0018	0.0525	-0.0018	1.0000	99.6442
Bosera	0.0072	0.0140	0.0000	0.1372	0.0156	0.0024	0.0155	0.0025	0.9979	99.9686
E Fund	0.0033	0.0029	0.0000	0.0151	0.0043	-0.0009	0.0028	-0.0008	0.9963	99.9990
September 2015-March 2016										
GuoTai	0.0021	0.0026	0.0000	0.0249	0.0033	-0.0004	0.0033	-0.0003	0.9998	99.9986
HuaAn	0.0272	0.0312	0.0000	0.1453	0.0414	-0.0014	0.0415	-0.0011	0.9961	99.7817
Bosera	0.0066	0.0115	0.0000	0.083	0.0132	0.0013	0.0132	0.0011	1.0017	99.9781
E Fund	0.0029	0.0026	0.0000	0.0125	0.0039	-0.0005	0.0027	-0.0002	0.9969	99.9990

Note: Tracking errors are expressed as percentages.

Table 3. The determinants of tracking error

	α	β_1	β_2	β_3
TE_1				
Guotai	-0.0024	-0.0000*	0.0007*	0.0020
Huaan	0.2030**	-0.0000	0.0000	-0.0743**
Bosera	-0.2735**	0.0000	-0.0001	0.1221**
E Fund	0.0002	0.0000	-0.0002	0.0012
TE_2				
Guotai	0.0014	-0.0000	0.0000	0.0004%
Huaan	0.0132	0.0000	-0.0021	0.0068
Bosera	-0.0524**	0.0000**	-0.0734**	0.0249**
E Fund	0.0000	-0.0011	0.0036	-0.0052
TE_3				
Guotai	0.0019	-0.0000*	0.0005*	0.0021
Huaan	0.2010**	0.0001	0.0000	-0.0744**
Bosera	-0.1235**	0.0000	-0.0001	0.0921**
E Fund	0.0002	0.0001	-0.0000	0.0011

Note: Coefficients are expressed as percentages; *t*-statistics are reported in parentheses; *indicates significance at 10% level, while **indicates significance at 5% or better.