ETP PERFORMANCE MEASUREMENT: TRACKING PERFORMANCE

Understanding the different exchange-traded product (ETP) performance measures is challenging. Different issuers have different standards and offer little on interpreting the measures and why they are essential. They fall into two categories—tracking performance and risk-adjusted performance. Within the risk-adjusted performance category are single market price risk factor (or traditional) measures and multiple market price risk factor (multi-factor) measures. This section focuses on providing a framework for analyzing tracking performance.

Tracking error

The variable of interest in measuring tracking performance is called *tracking error (TE)* and is defined simply as:

$$TE_t = R_{F,t} - LR_{B,t},$$

where $R_{F,t}$ and $R_{B,t}$ are the daily fund and benchmark holding period returns, respectively, and L is the gearing ratio (often referred to as the "leverage ratio"). A gearing ratio of $L = 3$ (i.e., a "levered fund") has a benchmark return of 3 times the daily return of the benchmark. A gearing ratio of $L = -1$ has a benchmark return of –1 times the benchmark return (i.e., an "inverse fund" but not a levered fund).

Sources of tracking error

Tracking error arises from many sources.

1) *Total expense ratio*. The total expense ratio includes operating costs and a management fee. If we hold other factors constant, the expense ratio causes the average mean tracking error to be less than 0.
2) *Transaction and rebalancing costs*. These are costs incurred when securities enter or leave the benchmark index. For passive funds, these events should be rare; however, they do occur. For example, a corporate event such as a cash dividend requires reinvestment. The index provider may change the composition of the benchmark, in which case the fund needs to rebalance its portfolio holdings. Also, rebalancing is routine for some benchmarks (e.g., equal-weighted indexes). The rebalancing costs, like the expense ratio, contribute to a negative mean tracking error.
3) *Sampling and optimization*. The fund manager often holds some securities in the benchmark index because it is impractical, infeasible, or too costly. Certain bond indexes, for example, contain thousands of securities, many of which are illiquid. In these instances, the fund will take a subset or sample of the index securities, weighting them to maximize the correlation between the fund and benchmark

returns or minimize its *TE*. SPY and VOO fully replicate the S&P 500, while IVV may sample/optimize. Sampling and optimization techniques have no predictable impact on tracking errors. Selecting a subset of securities may produce winners or losers relative to the complete index and, therefore, has no predictable effect on the mean *TE*.

4) *Cash drag*. Often, the benchmark index pays dividends daily, while the ETF pays quarterly. Because the daily benchmark dividends are too small to warrant immediate cost-efficient reinvestment in the benchmark securities, they may be held temporarily in cash, creating a cash drag. Cash drag will usually negatively influence tracking error since risky securities have higher expected returns than cash. Occasionally, the benchmark may perform poorly, and cash will have a higher return.

5) *Timing*. When an index provider changes the composition of a benchmark, changes are instant. An ETP, on the other hand, must execute its trades more slowly through time. During this interval of time, market prices may move. The effect on the tracking error is unpredictable.

6) *Securities lending*. ETFs may lend securities in their portfolios, thereby generating securities lending income and positive tracking error.[1]

Analyzing tracking error

We specify a series of statistical measures used in tracking error analysis with the daily tracking errors in hand. The *tracking difference* (*TD*) is the average of the difference between the fund's return and its benchmark return, that is,

$$TD = \frac{\sum_{t=1}^{T}\left(R_{F,t} - LR_{B,T}\right)}{T}, \tag{1}$$

where *T* is the number of days in the observation period. This performance measurement is not risk-adjusted because whatever price risk factors affect the fund return also affect the benchmark return and vice versa. The fund's objective is to match the returns of the benchmark, and the *TD* shows the degree to which it has. The *standard deviation of tracking error* measures the degree to which the TE deviates from its mean,

$$\sigma_{TE} = \sqrt{\frac{\sum_{t=1}^{T}\left(TE_t - TD\right)^2}{T-1}}, \tag{2}$$

and the standard error of the tracking difference is $\sigma_{TD} = \frac{\sigma_{TE}}{\sqrt{T}}$.

1

The *risk-adjusted tracking difference* is

$$RATD = \frac{TD}{\sigma_{TD}}. \tag{3}$$

RATD measures risk-adjusted performance and provides a *t*-test statistic for the null hypothesis that the TD equals 0.

Popular misconception

Industry practice often measures tracking error performance as the variability of the fund return and the benchmark index return, that is,

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}\left(R_{F,t} - LR_{B,T}\right)^2}{T}}. \tag{4}$$

Expression (4) is <u>not</u> a performance measure. It is nothing more than the *root mean squared error (RMSE)* definition in statistics. The *RMSE* may be re-written as

$$RMSE \approx \sqrt{\sigma_{TE}^2 + TD^2}. \tag{5}$$

In other words, a fund manager is penalized for σ_{TE}^2 the variance of the tracking error (which is fine) <u>and</u> for positive *TD* (outperformance) or negative *TD* (underperformance) in an equal manner (which is silly). Other industry participants refer to the formula for *RMSE* as the *standard deviation* of the tracking errors. That is technically not true.

<u>Illustration 1</u>: Compare tracking performance of SPY, IVV, and VOO.

The Excel file, ***S&P 500 ETF tracking performance.xlsx****, contains ten years of daily total return indexes for SPY (State Street), IVV (BlackRock), and VOO (Vanguard) ETFs on the S&P 500 index portfolio. Evaluate their tracking performances from the perspective of someone who wants to buy and hold a well-diversified stock portfolio.*

The results are not surprising in many ways. Recall that when we discussed the most popular ETPs in the market, all three appeared in the top 5 in terms of $AUM. In reviewing their current market activity, we found that even though SPY had the highest expense ratio at 0.0945% (relative to IVV and VOO at 0.03%), it had the highest $AUM and $TV. Based on the length of the average holding period of the funds, we concluded that SPY is a short-speculation and risk management tool, and IVV and VOO are long-term investment tools.

With these understandings in mind, we turn to the tracking error results. The tracking difference (i.e., mean tracking error) is lowest at -0.000005. IVV and VOO

had tracking differences of -0.000001. If we multiply these values by the number of trading days in a year, 252, SPY is -0.12%, while IVV and VOO are -0.03%. These values correspond closely to their stated expense ratios.

Interestingly, the standard deviation of the tracking error is higher for SPY than for IVV and VOO. One possible explanation for this behavior is that SPY has greater trading activity than the other two. The range of tracking errors also supports this conclusion. Based on the results, IVV and VOO appear to perform equally well mimicking the performance of the S&P 500 index portfolio.

Sample period	
Begins	20121231
Ends	20221230
No. of days	3,651
No. of years	10.00

Tracking performance			
Description	SPY	IVV	VOO
No. of daily returns (n)	2,518	2,518	2,518
Tracking difference (TD)	-0.000005	-0.000001	-0.000001
Standard deviation of TE	0.000562	0.000511	0.000511
Standard error of TD	0.000011	0.000010	0.000010
Risk-adjusted tracking difference ($RATD$)	-0.4094	-0.1216	-0.1197
Prob $TD = 0$	0.6823	0.9032	0.9048
Autocorrelation	-0.5767	-0.4642	-0.4190
Minimum	-0.0077	-0.0041	-0.0053
Median	0.0000	0.0000	0.0000
Maximum	0.0104	0.0042	0.0055
Root mean square tracking error ($RMSE$)	0.0006	0.0005	0.0005
Annualized TD (252 days)	-0.0012	-0.0003	-0.0003
Correlation between ETP and benchmark	0.9988	0.9989	0.9989