

# Bitcoin ETFs: Measuring the performance of this new market niche

Alessio Brini      Jimmie Lenz

An exchange traded fund (ETF) is a financial product that attempts to match the performance of a benchmark, which might be an index like the S&P 500, or a particular commodity. A key characteristic of the ETFs is that they can be traded on stock exchanges, just as individual equities are traded. This is particularly valuable to retail investors, for whom obtaining exposure to equity indices or commodities is otherwise only available through mutual funds. ETFs provide cheaper and easier access to a huge range of assets and indices, and hence have become very popular.

Recent additions to the ETF ecosystem have included products that attempt to track the price of Bitcoin (BTC), the most valuable cryptocurrency among digital assets. ProShares Bitcoin Strategy ETF (BITO) began trading on October 18th last year, followed by Valkyrie (BTF) and VanEck (XBTF). Managers of these products use the futures markets to track the price of Bitcoin, yet there are no ETFs that tie directly to Bitcoin's spot price.

The use of futures to track BTC may introduce a difference between the return of the ETF and that of its underlying benchmark. The standard deviation of this difference – the so-called “tracking error” – is the typical measure of performance for an ETF relative to its index and can be of considerable magnitude, which is troubling, since these ETFs have already attracted billions in investment dollars.<sup>1</sup> Furthermore, tracking error alone may be insufficient as a measure of ETF performance. The less commonly used “tracking difference” metric provides additional insight into BTC-benchmarked ETF that should also be taken into consideration. In this article, we disentangle the various components of tracking error, and explain the use of the additional tracking difference metric to provide an enhanced framework for evaluating the Bitcoin ETFs performance.

---

<sup>1</sup>ProShares reached over \$1 billion in its first day of trading

# 1 Part I: How well do the Bitcoin ETFs track the benchmark?

Tracking error is typically measured as the standard deviation of daily return differences between the ETF and its benchmark, thus:

$$TE = \sqrt{\frac{\sum_i^N (r_i^{ETF} - r_i^{\text{benchmark}})^2}{N - 1}}$$

where N is the number of observations used to compute the quantity.<sup>2</sup> Let's review the causes of these discrepancies:

1. *The investment in futures rather than the underlying Bitcoin.* When the ETF's exposure to the underlying asset is obtained via futures rather than spot BTC, trading limits may prevent the manager from investing 100% of the fund in the futures. Moreover, since futures are traded on margin and do not require the full value of the derivative – the so-called notional value – to be invested, the remaining funds are invested in highly liquid money market instruments and disinvested when needed to be added to the margin account. This practice helps in earning a short-term rate of return, which contributes to the total return of the funds. However, it can generate performance “slippage” between the ETF and BTC.<sup>3</sup>
2. *Exchange limits on futures positions.* The Chicago Mercantile Exchange (CME), on which Bitcoin futures are traded, imposes maximum futures holding positions on ETF issuers. If demand for the Bitcoin ETFs is sufficiently high, the fund managers may be forced to purchase longer term futures contracts in addition to the front month contract. This introduces “basis risk” (pricing differences between the different futures contracts).
3. *Transitions between futures contracts.* Futures are derivative products with a fixed maturity. To continuously track the benchmark, a manager of a futures-linked ETF must periodically transition the fund from the expiring (“front month”) contract into a futures contract with a later expiration date. This futures “roll” has an associated “roll yield”, which may be positive or negative depending on the relative prices of front month and longer-dated futures contracts (Schwager, 2017). Note that the roll yield is uncorrelated with the return performance of the underlying asset (BTC).
4. *Different trading hours.* Cryptocurrency exchanges allow investors to trade BTC around the clock, whereas ETFs products are only traded

---

<sup>2</sup>Note that, for comparison purposes across ETFs, this daily tracking error measure is typically annualized, using the standard technique: multiplying the daily volatility by the standard deviation of the number of trading days per year: thus, the annual tracking error is estimated as  $TE * \sqrt{252}$ .

<sup>3</sup>Investors can observe this difference in the ETF's fact sheet, which reports the holding composition. (See Rompotis (2016))

during standard market hours (Lin and Mackintosh, 2010). This misalignment in trading hours can produce a discrepancy between the ETF and the benchmark return. In the case of a BTC futures-linked ETF, tracking error is calculated through closing prices, while the Bitcoin is continuously traded. Therefore, it can happen that a sudden shock in Bitcoin price volatility is not immediately reflected by the ETF quoted prices until the market in which the product is exchanged reopens.

To summarize: returns on Bitcoin ETFs have multiple components: spot market return on the underlying Bitcoin, money market returns (or basis risk) due to the investments of the remaining cash after trading futures on margin and either positive or negative return associated with roll yield. The latter metrics represent the sources of tracking error. A cautious investor should also monitor the amount of the expenses for managing the fund, which represent fixed costs to pay when holding an ETF product. Usually, they are also relatively high for this category of financial instrument: 0.95% yearly for BITO, 0.83% for USO (crude oil), 0.4% for GLD (gold), as reported in each fact sheet provided by the ETF issuers. If the total expense ratio (TER) of the fund is fixed over time, management expenses does not constitute a source of tracking error, while they just contribute to increase the difference in return between the ETF and its benchmark (Charteris and McCullough, 2020). Instead, when the TER varies, management expenses are included as an additional source of tracking error.

## 2 Part 2: Is Tracking Error the best metric for evaluating ETF performance?

Tracking error, as explained above, is the standard tool used by practitioners and academics alike for measuring the performance of an ETF. However, it may not be the ideal tool from an investor perspective, as we now discuss.

1. *Tracking Difference: an alternative measure.* Fund fact sheets provided by ETF issuers often report an alternative measure: Tracking Difference (Charteris and McCullough, 2020). This is computed as the difference in cumulative returns of the ETF versus its benchmark and is thus a better measure of a fund's relative return performance. Tracking error, on the other hand, measures fluctuations in the return difference over time, thus providing an estimate of the relative risk of the ETF with respect to its benchmark. Thus, both of these measures should be examined in determining ETF performance, because, as with any investment, investors need to evaluate both return and risk. A good ETF should have both a small performance gap (tracking difference), and low fluctuations of this gap (tracking error). Only if both of these measures are small, we can confirm that the fund tracks the benchmark in a consistent manner.
2. *Return horizons and tracking error.* Another potential source of concern with the tracking error measure is the return horizon used for compu-

tation (Lin and Mackintosh, 2010). It has been shown that increasing the horizon from daily to annual returns leads to a significant drop in the annualized measure of tracking error. This occurs because short-term mispricings of an ETF with respect to its benchmark tend to mean revert over time. Tracking error measured over longer horizons do not account for this mean reversion. Tracking error can thus overestimate the true riskiness of the fund and needs to be considered within the context of the investor. Long-term investors will be less concerned about daily deviations from the benchmark since the mean reversion renders them irrelevant over a year or more. On the other hand, short-term investors and traders should pay close attention to the daily deviation patterns.

3. *Net Asset Value as evaluation metric.* As a measure of performance, tracking error requires a certain attention and a bit of context when used for doing this type of analysis. The optimization process for the ETF construction makes the standard deviation of return differences a misleading measure of the real performance of the fund. The different portfolio construction is properly evaluated by looking at the total value of the components themselves through a measure called Net Asset Value (NAV). Even if NAV can suffer intraday fluctuations, typically caused by market hour differences, it represents a more realistic measure of the value of the fund rather than its quoted price. NAV allows the investor to directly assess how much the fund is worth and should be compared to the value of the benchmark over the investment horizon. Carrying out an ETF evaluation on quoted prices instead of NAV represents a frequent misuse of performance measures such as the tracking error or the tracking difference. For instance, the level of liquidity of an ETF highly affects if it is traded at a premium or at a discount price with respect to the NAV (Rompotis, 2016). Hence, using the former to compute the returns of an ETF can be misleading because of this mispricing effect.

Futures-linked ETFs on BTC are a protected investment with all the underlying guarantees of a product traded on a stock exchange that is backed by standardized and regulated futures contracts on the CME. The investor does not need to bear the risk of holding BTC on his own through a private DeFi wallet. However, the advantage of obtaining BTC exposure via ETFs on futures contracts should be weighed against both tracking error and tracking difference of the ETFs vs spot BTC.

## References

- Charteris, A. and McCullough, K. (2020). Tracking error vs tracking difference: Does it matter? *Investment Analysts Journal*, 49(3):269–287.
- Lin, V. and Mackintosh, P. (2010). Etf mythbuster: Tracking down the truth. *The Journal of Index Investing*, 1(1):95–106.

Rompotis, G. G. (2016). Physical versus futures-based replication: The case of commodity etfs. *The Journal of Index Investing*, 7(2):16-37.

Schwager, J. D. (2017). *A Complete Guide to the Futures Market: Technical Analysis, Trading Systems, Fundamental Analysis, Options, Spreads, and Trading Principles*. John Wiley & Sons.