Humanities and Social Sciences Letters
2021 Vol. 9, No. 4, pp. 389-402.
ISSN(e): 2312-4318
ISSN(p): 2312-5659
DOI: 10.18488/journal.73.2021.94.389.402

DOES AN EXCHANGE-TRADED FUND CONVERGE TO ITS BENCHMARK IN THE LONG RUN? EVIDENCE FROM ISHARES MSCI IN ASIA-PACIFIC COUNTRIES

Kok-Leong Yap[1]
Wee-Yeap Lau[2+]
Izlin Ismail[3]

[1,3]*Faculty of Business and Accountancy, University of Malaya, Malaysia.*
[1]*Email: wallaceyap@hotmail.com*
[3]*Email: izlin@um.edu.my*
[2]*Faculty of Economics and Administration, University of Malaya, Malaysia.*
[2]*Email: wylau@um.edu.my*

(+ Corresponding author)

Article History
Received: 3 August 2021
Revised: 18 October 2021
Accepted: 5 November 2021
Published: 7 December 2021

Keywords
Exchange-traded fund
Cointegration
Performance
Tracking error.

JEL Classification
G11; G12; G14.

ABSTRACT

This study examines the linkages between exchange-traded funds (ETFs) and their benchmark indices from 2013 to 2019 using iShares MSCI of ten Asia-Pacific countries. Our results show, first, that there is a long-run causality running from the benchmark index to ETFs. These findings imply that ETFs may replicate the performance of the benchmark index over the long run. Second, there is a unidirectional causal relationship from ETFs to the benchmark index in the short run, which indicates that benchmark index prices respond to the short-run changes in the ETF prices when new information is available. Third, there is a significant tracking error between ETFs and the benchmark index. This finding justifies the existence of stock selection and market timing abilities among the ETF managers. Lastly, fund managers add value to the ETFs and generate better than the market returns. This paper provides new evidence to support this new stylized fact of ETFs.

Contribution/Originality: This study contributes to the existing literature on whether ETFs replicate the performance of the benchmark index in the long run. The findings of long-run relationships between ETFs and benchmark index prices highlight the price discovery role of ETF, which could benefit benchmark index forecasting.

1. INTRODUCTION

Since their inception in the 1990s, exchange-traded funds (ETFs) have gained popularity among retail investors due to tax efficiency and lower costs than other index-linked products (Broman, 2016). Relevant issues that have been studied regarding ETFs include the price discovery process and information transmission of ETFs (Buckle, Chen, Guo, & Tong, 2018; Tse & Martinez, 2007), the diversification benefits of ETFs (Gad & Andrikopoulos, 2019; Neves, Fernandes, & Martins, 2019), the volatility forecasting of ETFs (Tseng, Lee, & Chen, 2015; Zhu, Luo, & Jin, 2019), and arbitrage effectiveness (Hilliard, 2014).

The nature of ETFs is designed to replicate the performance of the benchmark index. Theoretically, ETF prices should closely track the fundamentals of the benchmark index, and thus, the tracking error for an ETF should be zero. However, the tracking error still exists due to several factors, as Dorocáková (2017) mentioned. Another topical issue is pricing deviation and imbalance between ETFs and net asset values (NAV). According to finance theory, the trading price of an ETF should be equivalent to the NAV.

In practice, the price of an ETF is determined by the interaction of supply and demand. It does not depend on the NAV of the ETF at a specific date and time. Jares & Lavin (2004) noticed a deviation between ETFs and NAV prices. Marshall, Nguyen, & Visaltanachoti (2013) discovered that ETF prices deviate significantly from NAV during periods of market volatility.

In an earlier study by Ackert & Tian (2008), they claimed that lower liquidity of the ETF fund will lead to a more significant deviation between ETF and NAV prices. Piccotti (2018) provides evidence to confirm that the ETF tends to trade at a higher premium due to the liquidity benefits. Aber, Li, & Can (2009) discovered that ETFs are more likely to trade at a premium than a discount. The diversification benefits and barriers to foreign investment may entice investors to pay premiums for ETFs (Delcoure & Zhong, 2007).

Wong & Shum (2010) investigated 15 worldwide ETF performances across bullish and bearish markets from 1999 to 2007. Their findings showed that an ETF provides better returns in a bullish market than in a bearish market. Meanwhile, Shanmugham & Zabiulla (2012) found that an ETF's alpha return in a bearish market is higher than in a bullish market.

Blitz & Huij (2012) found that emerging market ETF tracking errors are substantially higher than developed markets. Likewise, Khan, Bacha, & Masih (2015) found that the ETFs in emerging markets are less efficient in tracking the benchmark index's performance and have more significant tracking errors than ETFs in developed markets.

On the other hand, Miziołek & Feder-Sempach (2019) discovered that the 14 ETFs listed on European stock exchanges that attempt to replicate the performance of the MSCI Emerging Markets Index are effectively managed, as the tracking error values are generally lower. Additionally, Buetow & Henderson (2012) found that the ETF closely tracks the benchmark index, especially for a benchmark index composed of liquid securities. Rompotis (2011) found that the tracking errors of ETFs persist in the short term. However, ETF and benchmark index deviation generate arbitrage opportunities (Marshall et al., 2013). Further, Alam (2013) revealed that both Islamic and conventional ETFs could outperform the benchmark index.

Broman (2016) provided further evidence of mispricing between ETF and NAV and suggested it is due to the ETF price. Meanwhile, Levy & Lieberman (2013) found that NAV returns mostly drive the ETF prices. Delcoure & Zhong (2007) noticed that price deviations between ETFs and NAVs are temporary phenomena, and the price deviations tend to converge to zero within a short period. Shanmugham & Zabiulla (2012) showed similar results that the price deviation disappears within three days due to the arbitrage mechanism. Petajisto (2017) observed that arbitrageurs actively use the ETF share creation and redemption process to trade against the mispricings between ETFs and NAVs.

Generally, the level of integration across ETF and benchmark index is essential to investors. Cointegration is frequently used to examine the degree of interdependence between two or more closely related financial markets (Dimpfl, 2014). The existence of cointegration between ETF and benchmark index implies that they are interdependent. If two markets move together over a long-run period, any price imbalances that arise in one market would be corrected in the long run (Sharma, Thuraisamy, Madyan, & Laila, 2019). In other words, it provides an opportunity for the investor to arbitrage the market.

Shin & Soydemir (2010) claimed that ETFs are less efficient in disseminating information and do not provide significant investment benefits. Elton, Gruber, Comer, & Li (2002), Gastineau (2004), and Blitz, Huij, & Swinkels (2012) showed similar results – that ETFs underperform relative to the benchmark index. Neves et al. (2019) further argued that the diversification benefit through an ETF is limited, especially in financial crises.

It is worth noting that the previous studies (Ackert & Tian, 2008; Alam, 2013; Blitz et al., 2012; Dorocáková, 2017; Elton et al., 2002; Levy & Lieberman, 2013; Marshall et al., 2013) primarily focused on the US or European markets and that little is known about the behaviour and the performance of ETFs in Asia. Yap, Lau, & Ismail (2021) compared the difference between Islamic and Conventional Exchange-Traded Funds. This study attempts to fill this gap by providing an in-depth analysis of the performance of ETFs and how well they replicate the benchmark index. In addition, the cointegration relationship between ETFs and the benchmark index is examined to understand the patterns of interaction between the two.

2. DATA AND METHODOLOGY

The first part of the methodology examines the cointegration between the ETF and benchmark index prices. The second part examines the annualized holding period return, Sharpe ratio, Jensen's alpha, tracking error, and premiums/discounts associated with the ETF.

This study uses ETFs from ten Asia-Pacific countries, namely Australia, China, Hong Kong, India, Japan, Korea, New Zealand, Singapore, Taiwan, and Thailand. All the ETFs are from iShares MSCI. Each ETF is used as a proxy for the respective ETF of the country. Table 1 describes the data and their sources.

It can be observed from the dataset that each ETF starts at different dates. The first trading day in 2013 is used as a means to standardize the study period. Hence, daily prices from January 2, 2013 to December 31, 2019 are used. All data were obtained from Datastream and Bloomberg.

Table-1. Summary of ETF ticker symbols, names, benchmark index, and the dates for the first available data point.

Symbol	ETF	Benchmark Index	Inception Date
EWA	iShares MSCI Australia	MSCI Australia Index	12/03/1996
MCHI	iShares MSCI China	MSCI China Index	29/03/2011
EWH	iShares MSCI Hong Kong	MSCI Hong Kong Index	12/03/1996
INDA	iShares MSCI India	MSCI India Index	02/02/2012
EWJ	iShares MSCI Japan	MSCI Japan Index	12/03/1996
EWY	iShares MSCI Korea	MSCI Korea 25/50 Index	09/05/2000
ENZL	iShares MSCI New Zealand	MSCI New Zealand IMI 25/50 Index	01/09/2010
EWS	iShares MSCI Singapore	MSCI Singapore 25/50 Index	12/03/1996
EWT	iShares MSCI Taiwan	MSCI Taiwan 25/50 Index	20/06/2000
THD	iShares MSCI Thailand	MSCI Thailand IMI 25/50 Index	26/03/2008

2.1. Cointegration Test

This study employs different techniques to investigate the relationship between the ETF and benchmark index prices. The first method is the Engle-Granger two-step cointegration methodology (Engle & Granger, 1987). This method is based on analyzing the stationarity of error terms, as per Equation 1 below:

$$y_t = \mu + \alpha_1 x_t + \varepsilon_t \qquad (1)$$

Where y_t and x_t represent two different market prices, ε_t is the residual, and μ is the constant. Next, the first difference of the residuals is regressed against the lagged term of residuals without a constant, as stated in Equation 2:

$$\Delta\, \hat{\varepsilon}_t = \alpha_1 \hat{\varepsilon}_{t-1} + \sum_{i=2}^{n} \alpha_i \,\Delta\, \hat{\varepsilon}_{t-i} + e_t \qquad (2)$$

Where $\hat{\varepsilon}_t$ is the estimated residual, and α is the estimated parameter. A hypothesis test on coefficient α_1 is conducted to determine whether cointegration exists. The null hypothesis of no cointegration is rejected if the t-statistic of the coefficient exceeds a critical value.

The second method used to examine the ETF and benchmark index relationship is the Gregory & Hansen (1996) cointegration test. This newer method can test for cointegration in the presence of breakpoints or structural changes. In particular, Gregory & Hansen (1996) have separated the structural break cointegration into three models. The first model (Equation 3) is a structural break in the intercept (level shift model), the second model

(Equation 4) is a structural break in the intercept affected by the trend (level shift with trend model), and the third model (Equation 5) is a structural break in the slope and intercept (regime shift model).

$$y_{1t} = \mu_1 + \mu_2 \phi_{t\tau} + \alpha y_{2t} + \varepsilon_t \quad (3)$$

$$y_{1t} = \mu_1 + \mu_2 \phi_{t\tau} + \beta t + \alpha y_{2t} + \varepsilon_t \quad (4)$$

$$y_{1t} = \mu_1 + \mu_2 \phi_{t\tau} + \alpha_1 y_{2t} + \alpha_2 y_{2t} \phi_{t\tau} + \varepsilon_t \quad (5)$$

Where μ_1 is the intercept before the shift, μ_2 is the intercept differential over the intercept at the time of the shift, cointegrating slope coefficients before the regime shift, change the slope coefficients, and dummy variable. If there is a structural break, $\phi_{t\tau}$ it is equal to 1, and 0 otherwise.

Similar to the Engle-Granger two-step approach, the null hypothesis of no cointegration between two variables is tested by stationarity in residuals using Equations 3 to 5. In this study, the method to test the stationarity in the residuals is based on the adjusted augmented Dickey-Fuller (ADF) test in Equation 6:

$$ADF_t^* = \inf_{\tau \in T} ADF(\tau) \quad (6)$$

2.2. Vector Error Correction Model (VECM)

When the vector autogression (VAR) series are cointegrated, the dynamic relations between ETF and benchmark index prices can be undertaken through a VECM. The VECM can be written as Equations 7 and 8:

$$\Delta y_t = \alpha_1 + \rho_1 ecm1_{t-1} + \sum_{i=0}^{k} \beta_i \Delta y_{t-1} + \sum_{i=0}^{k} \delta_i \Delta x_{t-1} + \varepsilon_t \quad (7)$$

$$\Delta x_t = \alpha_2 + \rho_2 ecm2_{t-1} + \sum_{i=0}^{k} \beta_i \Delta y_{t-1} + \sum_{i=0}^{k} \delta_i \Delta x_{t-1} + \varepsilon_t \quad (8)$$

Where β_i and δ_i are the short-run coefficients, and $ecm1_{t-1}$ and $ecm2_{t-1}$ are the error correction terms of how y and x react to deviations from long-run equilibrium.

2.3. Granger Causality Test

Next, the Granger causality test is conducted to determine if there is any influence between the ETF and benchmark index prices. Although cointegration implies that causality exists between the two series, it does not indicate the direction of the causal relationship.

According to Granger (1969), with two series x_t and y_t, which are integrated in the same order, if the past and present values of providing some useful information to forecast x_{t+1} at the time t, it can be said that y_t Granger-causes x_t. We use the Granger causality test to test the linear causality between ETF and benchmark index prices in the study. The equation for the Granger causality test is stated in Equations 9 and 10:

$$\Delta x_t = \mu + \sum_{i=1}^{p} \alpha_{1i} \Delta x_{t-i} + \sum_{i=1}^{p} \alpha_{2i} \Delta y_{t-i} + \varepsilon_{1t} \quad (9)$$

$$\Delta y_t = \mu + \sum_{i=1}^{p} \alpha_{1i} \Delta y_{t-i} + \sum_{i=1}^{p} \alpha_{2i} \Delta x_{t-i} + \varepsilon_{2t} \quad (10)$$

where p is the maximum lagged observation determined by the final prediction error (FPE), and the Akaike Information Criterion (AIC) ε is the prediction error. The null hypothesis of Granger-causality is that x_t does not Granger cause y_t, and vice versa. A rejection of the null hypothesis indicates the presence of Granger causality.

2.4. Performance of ETFs

This section investigates the performance of the ETF as a passive investment tool by calculating several measurements. The performance of an ETF is evaluated in terms of annualized holding period returns, Sharpe ratio, Jensen's alpha, tracking error, and the premium/discount.

2.4.1. Holding Period Return

The holding period return is computed by dividing capital gain and income by the initial value of an investment. The formula is expressed in Equation 11:

$$HPR = \frac{Income + P_{n+1} - P_n}{P_n} \quad (11)$$

Where P_n is the initial value and P_{n+1} is the ending value. Income is the dividend distributed during the period. Next, the holding period return is annualized through Equation 12:

$$Annualized\ holding\ period\ return = (1 + HPR)^{\frac{1}{n}} - 1 \quad (12)$$

Where n is the number of years.

2.4.2. Sharpe Ratio

For every investor, it is essential to understand the risks associated with a particular investment. Generally, riskier investments should compensate investors with higher returns, and safer investments should not have extreme price fluctuations. Thus, evaluating an investment's risk-adjusted return is of the utmost importance while making investment decisions. A lower standard deviation of returns will lead to a higher Sharpe ratio, whereas a higher standard deviation will lead to a lower Sharpe ratio. The formula of the Sharpe ratio is stated in Equation 13:

$$Sharpe\ ratio = \frac{R_p - R_f}{\sigma_p} \quad (13)$$

Where R_p is the portfolio return, R_f is the risk-free rate, and σ_p is the standard deviation of the portfolio return. The 10-year US treasury bill rate is taken as the proxy for the risk-free rate (Lemperiere et al., 2017; Sharpe, 1966).

2.4.3. Jensen's Alpha

Jensen (1968) developed the concept of Jensen's alpha to evaluate the performance of a portfolio. The idea behind this is to determine the abnormal return of a portfolio. The formula is expressed in Equation 14:

$$R_p - R_f = \alpha + \beta\left(R_m - R_f\right) + \varepsilon_t \qquad (14)$$

Where R_p is the portfolio return, R_f is the risk-free rate, α is alpha, β is beta, R_m is market return, and ε_t is the error term.

2.4.4. Tracking Error

The tracking error formula can be described by Equation 15:

$$Tracking\ error = \sqrt{\frac{\sum_{t=1}^{n}(R_p - R_b)^2}{n-1}} \qquad (15)$$

where R_p is the ETF returns, R_b is the benchmark returns, and n is the number of observations.

2.4.5. Percentage Price Deviations

This study focuses on examining the percentage price deviations of the ETF from its net asset values. The formula is stated in Equation 16:

$$Premium/Discount = \frac{P_t - NAV_t}{NAV_t} \times 100 \qquad (16)$$

Where P_t is the price, and NAV_t is the net asset value.

Table-2. Descriptive statistics of daily returns for the period from 2013 to 2019.

	Mean	Max.	Min.	Std. Dev.	Skewness	Kurtosis	Jarque–Bera
Panel A: Benchmark index							
Australia	0.0197	3.9437	-6.0460	1.0629	-0.2509	4.9359	293.48*
China	0.0323	6.0238	-6.3924	1.2044	-0.0855	5.5037	462.09*
Hong Kong	0.0315	5.5169	-5.1258	0.9546	-0.1978	5.8098	590.78*
India	0.0287	5.9956	-7.2067	1.0943	-0.2724	6.5178	929.77*
Japan	0.0392	7.7943	-6.1341	1.1238	-0.0667	6.8254	1075.06*
Korea	0.0141	5.2233	-5.2435	1.1142	-0.2083	4.4141	159.46*
New Zealand	0.0513	3.6729	-5.1411	0.8973	-0.2475	5.1195	347.59*
Singapore	0.0135	4.5553	-4.4548	0.8480	0.1377	5.8889	617.94*
Taiwan	0.0408	5.2779	-6.9558	0.9896	-0.3432	6.8129	1101.29*
Thailand	0.0198	6.2665	-6.1773	1.0476	-0.1316	7.5351	1514.20*
Panel B: ETF							
Australia	0.0183	5.9400	-7.8640	1.1157	-0.3842	6.1006	748.74*
China	0.0317	6.5755	-7.1573	1.3851	-0.1269	4.8653	260.04*
Hong Kong	0.0295	5.4301	-6.1665	1.0397	-0.1935	5.8608	611.49*
India	0.0275	5.6988	-6.3265	1.2991	-0.1703	4.7468	232.39*
Japan	0.0334	4.9608	-4.7336	1.0017	-0.2880	6.0272	696.74*
Korea	0.0120	5.4090	-6.4220	1.2058	-0.1955	4.2819	131.80*
New Zealand	0.0493	3.3475	-4.9354	0.9476	-0.1277	3.9619	72.6828*
Singapore	0.0118	4.4427	-5.1829	0.9363	-0.1423	5.4392	442.52*
Taiwan	0.0387	4.3657	-5.6298	1.0993	-0.2646	4.4253	169.60*
Thailand	0.0201	8.1987	-9.0168	1.2417	-0.0891	7.6364	1579.63*

Note: * denotes rejection at the 5% significance level.

3. RESULTS

Table 2 shows that New Zealand has the highest average daily benchmark index returns, while Singapore has the lowest. Likewise, New Zealand has the highest average daily ETF returns, while Hong Kong has the lowest.

The standard deviation of returns, which is a proxy for risk, provides valuable information on the dispersion of returns. Panel A of Table 2 shows that China has the highest standard deviation for the benchmark index returns, while Singapore has the lowest. Likewise, China has the highest standard deviation of ETF returns, while Singapore has the lowest, as displayed in Panel B of Table 2. In most cases, the standard deviation of an ETF is higher than the benchmark index, which implies that the ETF is generally more volatile than the benchmark index.

It can be noticed that all the daily returns of the ETF and benchmark index show negative skewness, which indicates that most of the data values are concentrated on the right of the mean. The negative skewness of data distribution indicates that an investor may expect frequent small gains and a few large losses in the ETF. Overall, the daily returns of the ETF and benchmark index are not normal with negative skewness and positive excess kurtosis. In general, all the ETF and benchmark index prices are highly correlated. The correlation values are above 0.99, implying that the ETF and benchmark index prices share the same direction (see Table 3). However, the magnitude of the move is unknown. Therefore, a cointegration test is employed to examine the cointegration relationship between ETF and benchmark index prices.

Table-3. Correlations of ETF and benchmark index prices.

Countries	Correlation
Australia	0.9959
China	0.9985
Hong Kong	0.9985
India	0.9960
Japan	0.9976
Korea	0.9963
New Zealand	0.9997
Singapore	0.9958
Taiwan	0.9982
Thailand	0.9987

Table-4. Results of the unit root tests.

	Level ADF	Level PP	1st Level ADF	1st Level PP
Benchmark index				
Australia	-1.6707	-1.7066	-41.5408*	-41.5387*
China	-1.1382	-1.0964	-37.9899*	-37.9033*
Hong Kong	-1.2435	-1.2812	-39.7907*	-39.7582*
India	-1.3869	-1.3691	-39.2526*	-39.2517*
Japan	-1.9183	-1.8774	-52.1902*	-53.5636*
Korea	-1.6164	-1.6471	-41.5556*	-41.5554*
New Zealand	-0.3780	-0.3945	-40.5441*	-40.5250*
Singapore	-1.3783	-1.6969	-39.3947*	-39.7379*
Taiwan	-0.5360	-0.5262	-41.2919*	-41.2869*
Thailand	-1.1416	-1.1794	-40.0437*	-40.0632*
ETF				
Australia	-1.8937	-1.8770	-42.4286*	-42.4823*
China	-1.2750	-1.2929	-40.7411*	-40.7417*
Hong Kong	-1.3693	-1.3994	-41.1226*	-41.1226*
India	-1.7210	-1.5694	-43.5191*	-43.7036*
Japan	-1.9049	-1.7923	-43.8688*	-44.0692*
Korea	-1.8607	-1.8475	-41.5979*	-41.5979*
New Zealand	-0.4219	-0.3860	-43.2478*	-43.2433*
Singapore	-1.7013	-1.9089	-41.5342*	-41.6743*
Taiwan	-0.7509	-0.7491	-41.9603*	-41.9630*
Thailand	-1.4747	-1.4055	-42.3489*	-42.3942*

Note: * denotes rejection at the 5% significance level.

Before the cointegration test, it is necessary to ensure that all the series' orders of integration are similar. The augmented Dickey-Fuller (ADF) and Phillips–Perron (PP) unit root tests are performed to check for the non-stationary condition of the series. The optimal lag length of the ADF and PP tests are determined by the Schwarz Information Criterion (SIC). The null hypothesis of the ADF and PP tests is that the series has a unit root.

Table 4 displays the results of the unit root tests. It shows that both the ADF and PP tests do not reject the null hypothesis of non-stationarity for all the series at the level form. However, the test is rejected when the series are first differenced. Since the series are integrated of order one, it is possible to conduct a cointegration analysis to investigate the relationship between ETF and benchmark index prices.

The cointegration relationship between ETF and benchmark index prices is first tested based on the Engle-Granger two-step approach, and the results are reported in Table 5. The ADF test results from the Engle-Granger cointegration test reject the null hypothesis of no cointegration between ETF and benchmark index prices for all the countries. This result implies that ETF prices have a cointegrated relationship with the benchmark index prices.

A cointegrating relationship implies that any deviations between the two prices will be corrected even when each price is integrated, and each may be very close to a random walk (Engle & Sarkar, 2006). In other words, any disequilibrium in prices resulting in one market would be corrected in the long run (Sharma et al., 2019).

Table-5. ADF test results from the Engle-Granger cointegration test residuals.

Dependent variable	Benchmark Index		ETF	
	tau-statistic	z-statistic	tau-statistic	z-statistic
Australia	-3.9958*	-33.4918*	-4.1062*	-35.3214*
China	-7.0522*	-112.2536*	-7.0914*	-113.8169*
Hong Kong	-5.1089*	-56.5244*	-5.5201*	-66.2145*
India	-4.7274*	-47.5411*	-4.8089*	-49.4448*
Japan	-11.1743*	-350.1464*	-11.1921*	-351.8673*
Korea	-4.8805*	-47.5430*	-5.9021*	-72.2963*
New Zealand	-7.1235*	-121.2674*	-7.1231*	-121.5015*
Singapore	-3.3958*	-22.8289*	-3.5082*	-24.3843*
Taiwan	-7.5106*	-134.0989*	-7.5432*	-135.9102*
Thailand	-11.2180*	-350.9662*	-11.2820*	-356.4964*

Note: * denotes rejection at the 5% significance level.

For robustness, the cointegration between the ETF and benchmark index prices is further examined using cointegration with the structural break proposed by Gregory & Hansen (1996). This model is used to check whether the structural break would affect the cointegration relationship. According to Ghosh & Kanjilal (2016), the shock of external events, such as economic crises, technological shocks, and policy changes, may affect the cointegration test results in long time-series data.

The Gregory and Hansen test can be used for single breakpoints and when the break is unknown. This study determines the break date by estimating the cointegration equations for all possible break dates. The null hypothesis of the Gregory–Hansen cointegration test is that there is no cointegration against the alternative hypothesis of a cointegrating relationship with a structural break. The null hypothesis of no cointegration is rejected if the test statistic obtained from the ADF test is more than the critical value.

Table 6 displays the results of the test for cointegration with a structural break. It can be observed that the ADF test statistic rejects the null hypothesis of no cointegration at the 5% level of significance for all the countries. These results confirm a cointegration relationship between the ETF and the benchmark index, as suggested by Engle–Granger's two-step approach.

Given the results of the Engle-Granger two-step approach and the Gregory and Hansen test, we can conclude that there is a long-run relationship between the ETF and benchmark index prices for all the countries under study. From an investment perspective, cointegration between the ETF and benchmark index prices provides an opportunity for traders to arbitrage the market.

Table-6. Results of the test for cointegration with a structural break.

Country	Dependent Variable	Cointegration Model	Breakpoint	ADF Test Statistic	5% Critical Value
Australia	Benchmark	CC	13 May 2015	-9.1008*	-4.61
	Index	CT	2 September 2015	-14.3744*	-4.99
		CS	7 July 2015	-9.6291*	-4.95
	ETF	CC	13 May 2015	-9.2084*	-4.61
		CT	2 September 2015	-14.4751*	-4.99
		CS	19 February 2016	-9.7278*	-4.95
China	Benchmark	CC	5 June 2015	-9.9665*	-4.61
	Index	CT	18 May 2015	-13.6057*	-4.99
		CS	5 June 2015	-10.1443*	-4.95
	ETF	CC	16 June 2015	-10.0332*	-4.61
		CT	18 May 2015	-13.7577*	-4.99
		CS	5 June 2015	-10.2136*	-4.95
Hong Kong	Benchmark	CC	9 July 2015	-7.7408*	-4.61
	Index	CT	8 June 2015	-11.6568*	-4.99
		CS	9 July 2015	-7.7489*	-4.95
	ETF	CC	9 July 2015	-7.7854*	-4.61
		CT	5 June 2015	-11.8133*	-4.99
		CS	9 July 2015	-7.7953*	-4.95
India	Benchmark	CC	21 August 2017	-6.3647*	-4.61
	Index	CT	26 July 2017	-9.7526*	-4.99
		CS	30 July 2015	-7.4202*	-4.95
	ETF	CC	21 August 2017	-6.4245*	-4.61
		CT	26 July 2017	-9.9003*	-4.99
		CS	30 July 2015	-7.5867*	-4.95
Japan	Benchmark	CC	14 December 2015	-10.8512*	-4.61
	Index	CT	17 June 2016	-20.7127*	-4.99
		CS	14 December 2015	-10.8902*	-4.95
	ETF	CC	16 December 2015	-10.8485*	-4.61
		CT	17 June 2016	-20.5676*	-4.99
		CS	16 December 2015	-10.9053*	-4.95
Korea	Benchmark	CC	20 October 2014	-8.7984*	-4.61
	Index	CT	7 November 2018	-15.3202*	-4.99
		CS	16 October 2014	-8.7380*	-4.95
	ETF	CC	20 October 2014	-8.8782*	-4.61
		CT	24 October 2018	-19.6100*	-4.99
		CS	20 October 2014	-8.8748*	-4.95
New Zealand	Benchmark	CC	8 June 2015	-8.6281*	-4.61
	Index	CT	7 August 2017	-11.6112*	-4.99
		CS	8 June 2015	-8.7377*	-4.95
	ETF	CC	8 June 2015	-8.6265*	-4.61
		CT	7 August 2017	-11.6175*	-4.99
		CS	8 June 2015	-8.7299*	-4.95
Singapore	Benchmark	CC	26 July 2017	-6.7083*	-4.61
	Index	CT	25 May 2018	-12.1118*	-4.99
		CS	26 July 2017	-6.7083*	-4.95
	ETF	CC	26 July 2017	-6.7331*	-4.61
		CT	25 May 2018	-12.1896*	-4.99
		CS	16 September 2016	-6.7497*	-4.95
Taiwan	Benchmark	CC	17 July 2015	-10.7050*	-4.61
	Index	CT	31 August 2016	-18.7036*	-4.99
		CS	17 July 2015	-10.7228*	-4.95
	ETF	CC	9 July 2015	-10.8496*	-4.61
		CT	31 August 2016	-18.7965*	-4.99
		CS	9 July 2015	-10.8910*	-4.95
Thailand	Benchmark	CC	2 May 2018	-10.2583*	-4.61
	Index	CT	13 February 2014	-11.2048*	-4.99
		CS	1 February 2016	-10.2884*	-4.95
	ETF	CC	2 May 2018	-10.3768*	-4.61
		CT	12 February 2014	-13.1615*	-4.99
		CS	2 February 2016	-10.4682*	-4.95

Note: * denotes rejection at the 5% significance level; CC, CT, and CS denote the level shift model, level shift with trend model, and regime shift model, respectively.

Table 7 displays the results of the vector error correction model. The error correction term reflects the speed of adjustment towards long-run equilibrium. It is vital to note that the sign must be negative and statistically significant if an error correction term is appropriate. Panel A of Table 7 reveals that the coefficients of the error correction terms are statistically insignificant at the 5% significance level. These results suggest that there is no long-run causality running from ETF to benchmark index. Meanwhile, Panel B of Table 7 reveals that the coefficients of the error correction terms are statistically significant and negative, implying a long-run causality running from benchmark index to ETF. In other words, the ETF may replicate the performance of the benchmark index over the long run.

In addition, the coefficient of the error correction term indicates the speed of adjustment to equilibrium. It can be observed that Thailand has the highest speed of adjustment to equilibrium, with 46.53% of disequilibrium being corrected in a day. Meanwhile, India has the lowest speed of adjustment to equilibrium, with 6.61% being corrected in a day. The massive difference between the speed of adjustment to equilibrium may be partly due to the efficiency levels in different markets.

Table-7. Results of the vector error correction model.

Country	Error Correction Term	Standard Error	t-Statistic
Panel A: Dependent variable: Benchmark index			
Australia	0.0260	0.0266	0.9760
China	0.0366	0.0349	1.0488
Hong Kong	-0.0054	0.0280	-0.1942
India	0.0046	0.0203	0.2254
Japan	-0.0695	0.0456	-1.5266
Korea	-0.0248	0.0256	-0.9691
New Zealand	-0.0099	0.0553	-0.1786
Singapore	0.0183	0.0199	0.9225
Taiwan	-0.0159	0.0266	-0.5966
Thailand	0.0871	0.0763	1.1422
Panel B: Dependent variable: ETF			
Australia	-0.0825*	0.0357	-2.3105
China	-0.1446*	0.0481	-3.0085
Hong Kong	-0.0731*	0.0368	-1.9860
India	-0.0661*	0.0306	-2.1622
Japan	-0.1636*	0.0562	-2.9095
Korea	-0.0785*	0.0341	-2.3000
New Zealand	-0.2028*	0.0648	-3.1315
Singapore	-0.0726*	0.0270	-2.6849
Taiwan	-0.1194*	0.0373	-3.1994
Thailand	-0.4653*	0.1007	-4.6202

Note: * denotes rejection at the 5% significance level.

Subsequently, Granger causality tests are performed to explore the causal relationship between ETF and benchmark index prices in the short run. Table 8 reports the results of the Granger causality test. The null hypothesis of no Granger causality from ETF prices to benchmark index prices is rejected at the 5% significance level for all the countries, implying that changes in ETF prices lead the changes in benchmark index prices in the short run.

Meanwhile, the Granger causality tests show mixed results for the null hypothesis of no Granger causality from benchmark index prices to ETF prices. The test statistics suggest that the null hypothesis of no Granger causality from benchmark index prices to ETF prices is rejected at the 5% significance level for India, Japan, Singapore, Taiwan, and Thailand. However, the test fails to reject the null hypothesis for Australia, China, Hong Kong, Korea, and New Zealand. We can deduce a unidirectional causal relationship from the ETF to the benchmark index in the short run. However, there is a lack of evidence of a unidirectional causal relationship from the benchmark index to the ETF in the short run.

Table-8. The result of the Granger causality test.

Country	Benchmark Index → ETF	ETF → Benchmark Index
Australia	33.4722	480.5704*
China	16.4579	278.8363*
Hong Kong	31.7361	302.6669*
India	33.7747*	347.8429*
Japan	53.1477*	215.4218*
Korea	17.7010	373.5010*
New Zealand	8.6167	107.4867*
Singapore	32.7726*	430.8220*
Taiwan	36.1688*	338.9747*
Thailand	44.2943*	102.9048*

Note: "→" denotes the unidirectional causality relation; * denotes rejection at the 5% significance level.

Finally, we examine ETF's performance in holding period returns, Jensen's alpha, the Sharpe ratio, and tracking error. Table 9 shows that the annualized holding period return for all the countries is positive. The highest annualized holding period return is achieved by Hong Kong, while the second-highest performance is in New Zealand.

Table 9 also shows that the Sharpe ratio for all the ETFs is positive. Comparing the Sharpe ratios, New Zealand has the highest at 0.7505, while Taiwan has the second-highest at 0.4927. A higher Sharpe ratio indicates that the return per unit of risk is higher than others.

Table-9. Performance of ETFs.

Country	Annualized Holding Period Return	Jensen's Alpha	Sharpe Ratio	Tracking Error	Price Deviation
Australia	3.09%	0.0032	0.1960	0.8981	0.0426%
China	5.70%	0.0049	0.3106	0.9815	0.0423%
Hong Kong	17.76%	0.0036	0.3802	0.7429	-0.0356%
India	4.92%	0.0038	0.2810	0.9521	0.1643%
Japan	7.40%	0.0170	0.4571	1.1827	0.0762%
Korea	1.19%	0.0012	0.0978	1.0407	-0.0854%
New Zealand	11.94%	0.0083	0.7505	0.6667	-0.0201%
Singapore	1.88%	9.95E-05	0.1229	0.6665	0.0218%
Taiwan	8.55%	0.0114	0.4927	0.9802	-0.0259%
Thailand	3.17%	0.0031	0.1993	0.9151	0.0168%

Note: * denotes rejection at the 5% significance level.

Also reported in Table 9 is that Japan has the biggest tracking error, while Singapore has the smallest tracking error. Tracking error reflects the ETF's ability to track the return of the underlying benchmark closely. The smaller the tracking error, the better the replication strategy of the ETF.

The premium or discount of an ETF is the deviation of the ETF's price from its underlying net asset value (NAV). If the price of the ETF is above the NAV, the ETF is deemed to be trading at a premium. According to the law of one price, the ETF price should be in line with the NAV; any deviation between the prices indicates the presence of an inefficiency (Charteris, Chau, Gavriilidis, & Kallinterakis, 2014).

The existence of market inefficiency implies that arbitrageurs can take opposite positions in the ETF and underlying portfolio, unwind the positions using the ETF's creation or redemption process to arbitrage the price discrepancy. Jares & Lavin (2004) and Petajisto (2017) claim that a simple trading rule based on this mispricing produces impressive gross returns.

Table 9 also provides summary statistics on the distribution of the positive and negative percentage price deviations of the ETF and its net asset values. All ETFs exhibit positive premiums except Hong Kong, Korea, New Zealand, and Taiwan. Consistent with previous studies (Aber et al., 2009; Almudhaf & Alhashel, 2020), our results indicate that the ETFs are more likely to trade at a premium than at a discount. This phenomenon creates an arbitrage opportunity for arbitrageurs to buy the underlying securities and sell the ETF shares to profit from the mispricing.

4. CONCLUSION

This study investigates the cointegration relationship between the ETF and benchmark index prices. In addition, this study examines the performance of an ETF in terms of annualized holding period return, risk-adjusted return, tracking error, and the premium/discount associated with the ETF.

The Engle-Granger cointegration test and Gregory and Hansen cointegration test confirm a relationship between the ETF and benchmark index. The presence of cointegration indicates the existence of a long-term equilibrium relationship between the two.

Since the ETF and benchmark index prices are cointegrated, the vector error correction model (VECM) is employed. The results reveal a long-run causality running from the benchmark index to the ETF, which suggests that the ETF may be an appropriate substitute for directly investing in the stock market because it may replicate the performance of the ETF benchmark index over the long run.

A unidirectional causality from the ETF to the benchmark index indicates that benchmark index prices respond to the short-run changes in the ETF prices when new information is available. A detailed analysis of the Granger causality test provides information to ETF participants on causal relationships among markets that can help determine investment strategies.

This study also investigates the performance of ETFs as passive investment tools by calculating various measurements. In general, our findings reveal that ETFs show positive annualized holding returns and positive Sharpe ratios. The findings of tracking errors between ETFs and benchmark indices suggest stock selection and market timing abilities among the ETF fund managers. In other words, the ETF fund managers try to construct their investment portfolios to enhance the ETF returns. We can conclude that the ETF fund managers add value to the ETFs and generate better than the market returns, as suggested by a positive Jensen's alpha.

Our empirical results are practically useful for market participants to understand the pricing efficiency and performance of the ETFs and the relationship between them and benchmark indices. This finding provides further incentives to investors for considering ETFs as alternative instruments to investing in the stock market. The findings of long-run relationships between ETF and benchmark index prices highlight the price discovery role of ETFs, which could benefit benchmark index forecasting.

Funding: This study received no specific financial support.
Competing Interests: The authors declare that they have no competing interests.
Acknowledgement: All authors contributed equally to the conception and design of the study.

REFERENCES

Aber, J. W., Li, D., & Can, L. (2009). Price volatility and tracking ability of ETFs. *Journal of Asset Management, 10*(4), 210-221. Available at: https://doi.org/10.1057/jam.2009.13.

Ackert, L. F., & Tian, Y. S. (2008). Arbitrage, liquidity, and the valuation of exchange-traded funds. *Financial Markets, Institutions & Instruments, 17*(5), 331-362. Available at: https://doi.org/10.1111/j.1468-0416.2008.00144.x.

Alam, N. (2013). A comparative performance analysis of conventional and Islamic exchange-traded funds. *Journal of Asset Management, 14*(1), 27-36. Available at: https://doi.org/10.1057/jam.2012.23.

Almudhaf, F., & Alhashel, B. (2020). Pricing efficiency of Saudi exchange-traded funds (ETFs). *Journal of Islamic Accounting and Business Research, 11*(3), 793-809. Available at: https://doi.org/10.1108/jiabr-06-2017-0082.

Blitz, D., & Huij, J. (2012). Evaluating the performance of global emerging markets equity exchange-traded funds. *Emerging Markets Review, 13*(2), 149-158. Available at: https://doi.org/10.1016/j.ememar.2012.01.004.

Blitz, D., Huij, J., & Swinkels, L. (2012). The performance of European index funds and exchange-traded funds. *European Financial Management*, *18*(4), 649-662. Available at: https://doi.org/10.1111/j.1468-036x.2010.00550.x.

Broman, M. S. (2016). Liquidity, style investing and excess comovement of exchange-traded fund returns. *Journal of Financial Markets*, *30*, 27-53. Available at: https://doi.org/10.1016/j.finmar.2016.05.002.

Buckle, M., Chen, J., Guo, Q., & Tong, C. (2018). Do ETFs lead the price moves? Evidence from the major US markets. *International Review of Financial Analysis*, *58*, 91-103. Available at: https://doi.org/10.1016/j.irfa.2017.12.005.

Buetow, G. W., & Henderson, B. J. (2012). An empirical analysis of exchange-traded funds. *The Journal of Portfolio Management*, *38*(4), 112-127.

Charteris, A., Chau, F., Gavriilidis, K., & Kallinterakis, V. (2014). Premiums, discounts and feedback trading: Evidence from emerging markets' ETFs. *International Review of Financial Analysis*, *35*, 80-89. Available at: https://doi.org/10.1016/j.irfa.2014.07.010.

Delcoure, N., & Zhong, M. (2007). On the premiums of ishares. *Journal of Empirical Finance*, *14*(2), 168-195. Available at: https://doi.org/10.1016/j.jempfin.2005.12.004.

Dimpfl, T. (2014). A note on cointegration of international stock market indices. *International Review of Financial Analysis*, *33*, 10-16. Available at: https://doi.org/10.1016/j.irfa.2013.07.005.

Dorocáková, M. (2017). Comparison of ETF´ s performance related to the tracking error. *Journal of International Studies*, *10*(4), 154-165. Available at: https://doi.org/10.14254/2071-8330.2017/10-4/12.

Elton, E. J., Gruber, M. J., Comer, G., & Li, K. (2002). Spiders: Where are the bugs? *Journal of Business*, *75*(3), 453-472. Available at: https://doi.org/10.1086/339891.

Engle, R. E., & Granger, W. J. (1987). Cointegration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251-276. Available at: https://doi.org/10.2307/1913236.

Engle, R. F., & Sarkar, D. (2006). Premiums-discounts and exchange-traded funds. *The Journal of Derivatives*, *13*(4), 27-45. Available at: https://doi.org/10.3905/jod.2006.635418.

Gad, S., & Andrikopoulos, P. (2019). Diversification benefits of Shari'ah-compliant equity ETFs in emerging markets. *Pacific-Basin Finance Journal*, *53*, 133-144. Available at: https://doi.org/10.1016/j.pacfin.2018.10.009.

Gastineau, G. L. (2004). The benchmark index ETF performance problem. *The Journal of Portfolio Management*, *30*(2), 96-103. Available at: https://doi.org/10.3905/jpm.2004.319935.

Ghosh, S., & Kanjilal, K. (2016). Co-movement of international crude oil price and Indian stock market: Evidence from nonlinear cointegration tests. *Energy Economics*, *53*, 111-117. Available at: https://doi.org/10.1016/j.eneco.2014.11.002.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424-438. Available at: https://doi.org/10.2307/1912791.

Gregory, A. W., & Hansen, B. E. (1996). Residual-based tests for cointegration in models with regime shifts. *Journal of Econometrics*, *70*(1), 99-126. Available at: https://doi.org/10.1016/0304-4076(69)41685-7.

Hilliard, J. (2014). Premiums and discounts in ETFs: An analysis of the arbitrage mechanism in domestic and international funds. *Global Finance Journal*, *25*(2), 90-107. Available at: https://doi.org/10.1016/j.gfj.2014.06.001.

Jares, T. E., & Lavin, A. M. (2004). Japan and Hong Kong exchange-traded funds (ETFs): Discounts, returns, and trading strategies. *Journal of Financial Services Research*, *25*(1), 57-69. Available at: https://doi.org/10.1023/b:fina.0000008665.55707.ab.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of Finance*, *23*(2), 389-416. Available at: https://doi.org/10.1111/j.1540-6261.1968.tb00815.x.

Khan, A. P., Bacha, O. I., & Masih, A. M. M. (2015). Performance and trading characteristics of exchange-traded funds: Developed vs emerging markets. *Capital Markets Review*, *23*(1&2), 40-64.

Lemperiere, Y., Deremble, C., Nguyen, T.-T., Seager, P., Potters, M., & Bouchaud, J.-P. (2017). Risk premia: Asymmetric tail risks and excess returns. *Quantitative Finance*, *17*(1), 1-14. Available at: https://doi.org/10.1080/14697688.2016.1183035.

Levy, A., & Lieberman, O. (2013). Overreaction of country ETFs to US market returns Intraday vs. daily horizons and the role of synchronized trading. *Journal of Banking & Finance, 37*(5), 1412-1421. Available at: https://doi.org/10.1016/j.jbankfin.2012.03.024.

Marshall, B. R., Nguyen, N. H., & Visaltanachoti, N. (2013). ETF arbitrage: Intraday evidence. *Journal of Banking & Finance, 37*(9), 3486-3498. Available at: https://doi.org/10.1016/j.jbankfin.2013.05.014.

Miziołek, T., & Feder-Sempach, E. (2019). Tracking ability of exchange-traded funds. Evidence from emerging markets equity ETFs. *Bank and Credit, 50*(3), 221-248.

Neves, M. E. D., Fernandes, C. M., & Martins, P. C. (2019). Are ETFs good vehicles for diversification? New evidence for critical investment periods. *Borsa Istanbul Review, 19*(2), 149-157. Available at: https://doi.org/10.1016/j.bir.2019.01.002.

Petajisto, A. (2017). Inefficiencies in the pricing of exchange-traded funds. *Financial Analysts Journal, 73*(1), 24-54. Available at: https://doi.org/10.2469/faj.v73.n1.7.

Piccotti, L. R. (2018). ETF premiums and liquidity segmentation. *Financial Review, 53*(1), 117-152. Available at: https://doi.org/10.1111/fire.12148.

Rompotis, G. G. (2011). Predictable patterns in ETFs' return and tracking error. *Studies in Economics and Finance, 28*(1), 14-35. Available at: https://doi.org/10.1108/10867371111110534.

Shanmugham, R., & Zabiulla. (2012). Pricing efficiency of Nifty BeES in bullish and bearish markets. *Global Business Review, 13*(1), 109-121. Available at: https://doi.org/10.1177/097215091101300107.

Sharma, S. S., Thuraisamy, K., Madyan, M., & Laila, N. (2019). Evidence of price discovery on the Indonesian stock exchange. *Economic Modelling, 83*, 2-7. Available at: https://doi.org/10.1016/j.econmod.2019.09.005.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business, 39*(1), 119-138.

Shin, S., & Soydemir, G. (2010). Exchange-traded funds, persistence in tracking errors and information dissemination. *Journal of Multinational Financial Management, 20*(4-5), 214-234. Available at: https://doi.org/10.1016/j.mulfin.2010.07.005.

Tse, Y., & Martinez, V. (2007). Price discovery and informational efficiency of international iShares funds. *Global Finance Journal, 18*(1), 1-15. Available at: https://doi.org/10.1016/j.gfj.2007.02.001.

Tseng, T.-C., Lee, C.-C., & Chen, M.-P. (2015). Volatility forecast of country ETF: The sequential information arrival hypothesis. *Economic Modelling, 47*, 228-234. Available at: https://doi.org/10.1016/j.econmod.2015.02.031.

Wong, K. H., & Shum, W. C. (2010). Exchange-traded funds in bullish and bearish markets. *Applied Economics Letters, 17*(16), 1615-1624. Available at: https://doi.org/10.1080/13504850903085035.

Yap, K.-L., Lau, W.-Y., & Ismail, I. (2021). A comparative study between Islamic and conventional exchange-traded funds: Evidence from global market indices. *The Journal of Asian Finance, Economics, and Business, 8*(2), 725-735.

Zhu, F., Luo, X., & Jin, X. (2019). Predicting the volatility of the ishares china large-cap ETF: What is the role of the SSE 50 ETF? *Pacific-Basin Finance Journal, 57*, 101192. Available at: https://doi.org/10.1016/j.pacfin.2019.101192.