

# Multi-frequency Price Discovery in ETF Markets: Futures, Spot, and Net Asset Value Dynamics

Tzu-Pu Chang<sup>1</sup>, Jung-Che Tai<sup>1\*</sup> and Yi-Chi Lin<sup>1</sup>

## Abstract

This study empirically investigates the price discovery mechanism within the Taiwanese Exchange-Traded Fund (ETF) market by examining the dynamic lead-lag relationships between ETF futures, spot prices, and Net Asset Value (NAV). Focusing on three major ETFs (Yuanta Taiwan 50, Yuanta High Dividend, and Cathay Sustainable High Dividend) from August 2023 to August 2024, multi-frequency datasets are analyzed using Granger causality tests. Results reveal heterogeneous price discovery patterns across ETFs. For the Yuanta High Dividend ETF (0056) and Cathay Sustainable High Dividend ETF (00878), futures prices exhibit a dominant informational role, significantly leading spot prices and NAV, particularly in high-frequency (5-minute) data. In contrast, for the Yuanta Taiwan 50 ETF (0050), the price discovery process is more complex, with spot prices and NAV exerting a stronger causal influence on futures prices rather than the reverse. These findings suggest that the direction of information transmission varies with ETF characteristics and data frequency, highlighting the importance of multi-frequency analysis in understanding ETF market dynamics.

**JEL classification:** C01, C32, G00, G14.

**Keywords:** Exchange-Traded Funds (ETFs); Price Discovery; Futures; Granger Causality.

---

<sup>1</sup> Department of Finance, National Yunlin University of Science and Technology, Taiwan.

\*Corresponding author.

## 1. Introduction

The recent and rapid evolution of financial markets, along with a growing number of investment vehicles, has led to index investing becoming a key aspect of global capital markets. Exchange-Traded Funds (ETFs) integrate the intraday liquidity of individual stocks, with the diversification benefits of mutual funds, and so they have gained significant traction among both institutional and retail investors (Poutachidou & Koulis, 2025). Concomitant with the expansion of the ETF market, ETF futures have emerged as critical instruments for price discovery and risk management. Extensive academic literature underscores the dominant role of futures markets in the price discovery process. For instance, Tse (1995), demonstrated that futures price movements significantly led short-term adjustments in the spot price, and Alphonse (2000) found that over 95% of the price discovery process for the French CAC40 index occurred within the futures market. These papers highlight the informational efficiency and responsiveness of futures markets; nevertheless, there remains a lack of studies exploring ETF price discovery mechanisms, especially with respect to emerging markets.

The price formation of ETFs is a complex interplay of three distinct yet interconnected market segments: the futures market, the spot market, and the net asset value (NAV). Understanding the velocity and direction of information transmission is pivotal to understanding market efficiency and the price discovery mechanism. The futures market is generally perceived as more responsive, due to the lower transaction costs and superior liquidity, combined with the availability of financial leverage. This leading role in information dissemination is closely related to the rapid price adjustments. Despite the superior information that futures prices offer, statistical analyses of these markets may yield conflicting results when conducted using different time intervals. This makes the temporal efficiency of information transfer critical, particularly in a high-frequency trading (HFT) environment. Consequently, investigating the lead-lag relationships between ETF futures prices, spot prices, and NAV across various time scales has important theoretical and empirical implications regarding the underlying price discovery mechanism.

A review of previous literature reveals that there is no current consensus on whether ETFs and their derivatives possess an informational advantage, and empirical studies often depend on market maturity and specific characteristics of the underlying asset. Despite the general acceptance of futures market dominance, the literature still presents mixed findings regarding the informational role of the ETF itself, relative to its underlying spot price. For example, Kaur and Singh (2022), studying the Indian asset market, found that the spot price led ETF prices in the long run, suggesting limited ETF efficiency. Conversely, Atilgan et al. (2022), discovered that U.S. listed Emerging Market ETF returns had significant predictive power, demonstrating superior information integration. Furthermore, Kia et al. (2024) studied the highly complex Bitcoin market, and found that the information transmission path could vary widely for different market pairings. This compelling evidence underscores that prior literature has largely focused on analyses at a single frequency interval or within limited market dimensions, often neglecting to simultaneously compare the differences in information flow across a variety of time scales.

Given the aforementioned inconsistencies for the various markets and asset classes, this study focuses on three core questions, by which it can clarify the relative efficiency of the ETF market. First, it is necessary to examine whether the direction of information transmission for ETF futures, spot price, and NAV remains consistent for different data frequencies; and second, an analysis is required to determine whether futures markets exhibit informational leadership in the short term, thereby guiding the spot price and NAV. Finally, there must be an assessment of whether the stability and significance of the various price discovery mechanisms are affected by changes in data frequency. To address these questions, this research employs an empirical approach, utilizing data from three major Taiwanese ETFs that combines weekly, daily, and five-minute high-frequency data. The analysis begins with an Augmented Dickey-Fuller (ADF) test for stationarity, followed by a Vector Autoregression (VAR), to explore the dynamic relationships and determine any lead-lag structures. The analysis concludes with a Granger Causality Test, to verify the direction of information flow between the components.

The remainder of this paper is structured as follows: Section 2 is a literature review covering the theoretical and empirical foundations of ETF price discovery and futures market leadership; Section 3 details the data,

models, and methodology; Section 4 presents the empirical results and analysis; and Section 5 provides a conclusion and suggestions for further research.

## 2. Literature review

This section consolidates scholarly research on the price discovery mechanism, exploring the roles and functions of the futures and ETF markets with respect to information transmission and price formation. The review is structured into two main themes: futures price discovery and ETF price discovery. The futures market plays a pivotal role in the price discovery mechanism, and such research employs standardized econometric tests such as the ADF unit root test, VAR models, and the Granger Causality test, which can verify the futures market's role in information transmission and price adjustments, relative to the spot market. Tse (1995), analyzed the Japanese Nikkei Index, and employed an Error Correction Model (ECM), Univariate Time Series Method, and Vector Autoregression (VAR) to analyze price interaction. The results demonstrated that futures prices exhibited a short-term leading effect on the spot price, while the inverse effect was insignificant, suggesting the futures market holds a dominant position in information response and price adjustment. ECM exhibited the best predictive performance of the models the researchers tested. A study by Alphonse (2000) added further support to Tse (1995), using high-frequency data (every 30 seconds) for the French CAC40 index. Employing unit root tests, the Johansen Cointegration Test, ECM, and the Vector Moving Average Model (VMA), the study concluded that the futures market was the primary information source, with over 95% of the price discovery process occurring in the futures market. This result was consistent regardless of whether transaction prices or quotes were used, underscoring futures market dominance in price adjustment. Locally, Hsieh (2002) also found that futures prices held a dominant position in the Taiwanese stock index futures and spot markets, with evidence of information transmission links and co-integration between the two. Kavussanos et al. (2008) examined the lead-lag relationship in returns and volatility between the FTSE/ATHEX-20 and FTSE/ATHEX Mid-40 index futures and their cash markets in Greece. Using ADF and Phillips–Perron (PP) unit root tests, co-integration tests, VECM, Granger causality, and the VECM-GARCH-X model, the authors found that the futures market transmitted information faster than the cash market in terms of daily returns and volatility, particularly in the more liquid FTSE/ATHEX-20 market. The evidence that futures market volatility affects spot market volatility implies that the futures market plays a leading role in price discovery, and that it contributes to capital market efficiency and stability.

More recently, Hu et al. (2020) conducted an in-depth exploration of the relationship between Bitcoin spot and futures markets, focusing on price discovery and market guidance over time. This paper utilized advanced time-series methods, including Time-Varying Granger Causality, Time-Varying Cointegration, Time-Varying Information Share (IS), and the Dynamic Conditional Correlation GARCH (DCC-GARCH) model, analyzing futures markets on the CBOE and CME against the corresponding Bitcoin spot price. The study found that the futures market dominated the Bitcoin price discovery process during most periods, with futures prices Granger-causing spot prices, and importantly, that this causal relationship dynamically evolves over time. Similarly, Hao et al. (2019) compared the price discovery function of China's three major stock indices and their respective futures markets and investigated the impact of the 2015 trading restriction policy. Using 5-minute high-frequency data, VECM, Information Share (IS), and Component Share (CS) methods, they found the futures market dominated price discovery, a role that became more pronounced after the implementation of trade restrictions.

The performance and role of the ETF market in the price discovery mechanism is also an important research topic. Due to the dual influence of NAV and market supply/demand fluctuations on ETF prices, these analyses often rely on statistical or econometric methods to quantify the relative degree of influence. This investigation helps to clarify the ETF market's primary role in price discovery, and also provides a multi-faceted empirical method of evaluating market mechanisms.

Literature presents mixed findings with regard to the ETF's informational leadership role. Ho and Tseng (2007) found that the futures market in Taiwan showed limited trading activity during their study period (July 1, 2003, to June 28, 2005), leading to insignificant price discovery capabilities. Their study, using

daily close data and employing co-integration and VECM on the Taiwan 50 Index, futures, and ETF funds, concluded that the Taiwan 50 Index (Spot) ranked first in price discovery, followed by the ETF, with the futures ranking third. Kaur and Singh (2022) examined the information efficiency between Indian equity ETFs and their spot prices. Employing VECM, the Gonzalo and Granger Component Share Method, and the Hasbrouck Information Share Method, they found a long-term co-integration relationship where spot price changes led ETF price changes. The researchers suggested that the observed lack of ETF efficiency might be due to low market liquidity, limited public awareness of ETF products, and conservative promotion of passive funds by financial institutions.

In contrast, Atilgan et al. (2022), investigated the price discovery relationship between eighteen U.S.-listed Emerging Market equity ETFs and their spot prices. Using predictive regression models, and controlling for non-synchronous trading and autocorrelation to ensure accuracy, they found that ETF returns had significant predictive power over the underlying index's next-day returns, a finding that was especially pronounced during high volatility periods. The study validated the dominance of ETF prices over index prices via rolling window out-of-sample tests, and volatility-grouped analysis. While overall two-way information spillovers were weak, the ETF market showed better performance in price reaction speed and information integration, supporting the idea that ETFs are an informational leading indicator during market instability. Kia et al. (2024) explored the price discovery mechanism in the Bitcoin spot, futures, and ETF markets. Using minute-by-minute data and the Fractional Cointegrated Vector Autoregression Model (FCVAR) from the launch of the first U.S. Bitcoin ETF (October 19, 2021) to December 30, 2022, the FCVAR analysis showed co-integration between the spot and ETF markets, and between the spot and futures markets. Crucially, the spot market dominated in the spot-futures pairing, but the ETF market dominated the price discovery process in the spot-ETF pairing, underscoring the ETF's importance in overall Bitcoin price formation. Ji et al. (2024) investigated the dynamic dependence and volatility spillover effects among energy-related ETFs, with a focus on hedging effectiveness. Employing high-frequency data and time-varying models, their study demonstrated that the lead-lag relationships between ETF prices are highly dynamic and frequency-dependent, with short-term spillovers differing substantially from long-term interactions. These findings are consistent with the present study's motivation for adopting a multi-frequency analytical framework, further validating the importance of examining price discovery across different time scales.

In recent years, the ETF market has experienced explosive global growth. Generally, and particularly in Taiwan, this has been driven by factors related to cost structure, investor behavior, financial regulation, and new product offerings. ETFs offer lower management fees and greater transparency, compared to actively managed funds, establishing them as cost-efficient investment tools. SITCA (2025) data indicates that regulators in Taiwan are pushing to further reduce ETF management and custodial fees to enhance competitiveness. Furthermore, the widespread adoption of FinTech has lowered investment barriers, granting retail investors easy access through mobile apps and regular savings plans (RSPs). RSP account growth in 2024 shows that smaller volume, longer-term investment has become a major growth driver of the ETF market.

Additionally, institutional participation has significantly expanded the range of ETF product offerings. SITCA (2025), points to insurance companies becoming the primary source of funding for Taiwanese bond ETFs since 2019, reflecting a sustained institutional need for asset allocation and duration management. The surge of High Dividend ETFs is another key catalyst, with related data showing a substantial increase in Assets Under Management (AUM), and over 7.65 million beneficiaries in 2024; this accounts for over 70% of all Taiwanese ETF beneficiaries. The stable payouts and simple structure of these ETFs hold strong appeal for retail and retirement investors. Additionally, governmental policy supports such changes via promotion of the "Asia Asset Management Center", the opening of actively managed multi-asset ETFs, and by announcing plans to cross-list ETFs on multiple exchanges. Overall, the rapid growth of ETFs is a result of a synergy between cost efficiency, shifting investor demographics, expanding institutional demand, and policy support/product innovation.

This literature reveals that empirical results on price discovery are highly contingent upon the choice of econometric model, the specific characteristics of the asset, and the length of the analysis period. Past

studies often relied on a single statistical frequency, resulting in a less comprehensive view of the dynamic interaction among asset returns. Building upon this foundation, the current study provides a more specific discussion of financial market price discovery, by integrating the findings from various empirical models and research methodologies across different subject areas.

This research addresses existing research gaps by combining multiple time-series models to analyze the market from three perspectives: data stationarity, dynamic mutual influence, and leading market prediction. Crucially, the study introduces multiple data frequencies: weekly (for descriptive statistics), daily and 5-minute (for VAR and Granger causality analyses). In contrast with existing literature, this approach captures short-term and long-term price interactions. Furthermore, by considering ETFs with diverse product characteristics (0050, 0056, 00878) and including the NAV, this study also provides a comprehensive assessment of market price efficiency and potential arbitrage opportunities, thereby enriching the practical and theoretical understanding of the price discovery mechanism.

### **3. Methodology**

#### **3.1 Data**

This subsection describes the empirical models and sample data settings utilized for this study; where the aim is to analyze the price transmission direction and informational lead-lag relationships among ETF futures, spot prices, and NAV. The ETF market is selected as the focal subject for investigating the price discovery mechanism, due to its recent explosive growth in asset under management (AUM) and trading volume in the Taiwanese market. For empirical data, three Taiwanese ETFs are selected as representatives of their asset type, based on scale, stability, and liquidity, each being thought to reflect key consumer investment strategies: Yuanta Taiwan 50 (0050) (large-cap equities), Yuanta High Dividend (0056) (high-dividend strategy), and Cathay Sustainable High Dividend (00878) (sustainable and high-dividend). These selections are highly representative, and effectively capture the performance of different investment strategies within the Taiwanese ETF market. The research period is August 1, 2023, to August 1, 2024, and data is obtained from the XQ Global Winner System. The analysis utilizes the return series for three components of each ETF, including futures price, spot price, and NAV. Data analysis is conducted on daily and 5-minute frequencies for VAR and Granger causality tests, with weekly frequency data additionally included in the descriptive statistics, to provide a comprehensive view of both long and short-term market dynamics.

The econometric analysis is conducted using Stata software, and the empirical procedure commences by verifying the statistical properties of each time series. First, is the Augmented Dickey–Fuller (ADF) unit root test, which shows the stationarity of all return series variables (futures, spot, and NAV). Confirming stationarity is crucial for subsequent modeling, as it ensures that the returns are stable over time, and avoids the potential of spurious regression results. Following the confirmation of stationarity, the Vector Autoregression (VAR) model is constructed. This model is used for analyzing dynamic interactions and short-term influences among the three market return series. It enables a systematic investigation into the simultaneous co-movement of futures, spot prices, and NAV returns, and also reveals their dynamic linkages.

The final empirical step is the Granger Causality test within the VAR framework, which verifies the predictive capability of the variables. This test determines whether the historical information from the ETF futures (or any other component) formally leads the spot price or NAV. This is essential for establishing the direction of price transmission and the characteristics of information flow, thereby comprehensively revealing the price discovery mechanism in the Taiwanese ETF market. These three econometric methods are mutually complementary.

#### **3.2 Vector Autoregression Model (VAR)**

The Vector Autoregression (VAR) model, introduced by Sims (1980), serves as a fundamental tool in econometrics for capturing the dynamic interaction among multiple time series observations. The model operates under the assumption that all variables are endogenous, allowing for a systematic observation of

the dynamic impact of each variable on itself, and all other variables within the system. This framework is crucial for understanding the simultaneous co-movement and short-term linkages among the ETF futures, spot, and NAV return series.

The generalized VAR model of order  $p$ , denoted VAR( $p$ ), is formulated as follows:

$$F_t = \alpha + \sum_{i=1}^p \beta_i F_{t-p} + \sum_{i=1}^p \gamma_i S_{t-p} + \sum_{i=1}^p \delta_i NAV_{t-p} + \varepsilon_t \quad (3.1-1)$$

$$S_t = \alpha + \sum_{i=1}^p \gamma_i S_{t-p} + \sum_{i=1}^p \beta_i F_{t-p} + \sum_{i=1}^p \delta_i NAV_{t-p} + \varepsilon_t \quad (3.1-2)$$

$$NAV_t = \alpha + \sum_{i=1}^p \delta_i NAV_{t-p} + \sum_{i=1}^p \beta_i F_{t-p} + \sum_{i=1}^p \gamma_i S_{t-p} + \varepsilon_t \quad (3.1-3)$$

Where  $F_t$  = futures return;  $S_t$  = spot return;  $NAV_t$  = net asset value return;  $\alpha$  = intercept;  $p$  = lag order;  $\beta_i$  = futures lag coefficient;  $\gamma_i$  = spot lag coefficient;  $\delta_i$  = net asset value lag coefficient;  $\varepsilon_t$  = residual term. The VAR model can simultaneously consider the short-term interaction and dynamic relationship adjustment between futures, spot, and NAV returns, laying the foundation for subsequent Granger causal analysis.

### 3.3 Granger Causality Test

The Granger Causality Test, initially proposed by Granger (1969), is a fundamental econometric technique used to determine whether one time series variable possesses a predictive capability for another. The core principle asserts that if the past values of a variable  $X_t$  significantly improve the prediction of  $Y_t$  beyond what can be achieved using only the past values of  $Y$ , then  $X$  is said to Granger-cause  $Y$ . The Granger causality test formula is as follows:

1. Formula 3.2-1 does not demonstrate a causal relationship.

$$Y_t = \alpha_1 + \sum_{i=1}^p \beta_{1i} Y_{t-i} + \varepsilon_t \quad (3.2-1)$$

2. Formula 3.2-2 has a causal relationship.

$$Y_t = \alpha_2 + \sum_{i=1}^p \beta_{2i} Y_{t-i} + \sum_{j=1}^p \gamma_j X_{t-i} + \varepsilon_t \quad (3.2-2)$$

Where  $Y_t$  = dependent variable;  $X_t$  = explanatory variable (in this study model,  $Y_t$  can be represented as  $F_t$  = futures return;  $S_t$  = spot return;  $NAV_t$  = net asset value return),  $\alpha_1$ ,  $\alpha_2$  = intercept terms;  $\beta_{1i}$ ,  $\beta_{2i}$  = regression coefficients of the lagged dependent variable;  $p$  = lag order;  $\gamma_i$  = influence coefficient of  $X_t$  on  $Y_t$ ; and  $\varepsilon_t$  = residual term. The null hypothesis is  $H_0: \gamma_1 = \gamma_2 = \dots = \gamma_p = 0$ , that is,  $X_i$  has no predictive power for  $Y_t$ . If the test result rejects  $H_0$ , then  $X_i$  has a Granger causal relationship with  $Y_t$ .

## 4. Empirical result

Although weekly return data are included in the descriptive statistics to provide a comprehensive overview of return distributions across time horizons, the VAR and Granger causality analyses focus on daily and 5-minute frequencies. This is because the weekly sample size is substantially smaller, which may reduce statistical power and limit the reliability of VAR lag selection and Granger causality inference. This section presents the empirical results concerning the information transmission and price discovery relationships among the futures, spot price, and NAV returns for three selected ETFs: Yuanta Taiwan 50 (0050), Yuanta High Dividend (0056), and Cathay Sustainable High Dividend (00878). Consistent with the methodology established in Section 3, this empirical procedure includes the ADF unit root test to confirm variable stationarity, the VAR model to explore dynamic interactions, and the Granger Causality test to assess market lead-lag relationships.

### 4.1 Descriptive Statistics Summary

Table 1 presents the descriptive statistics for the returns of the three representative Taiwanese ETFs (0050, 0056, and 00878) across three price series: future, spot, and NAV. Overall, the mean returns for all three ETFs (weekly, daily, 5-minute interval), are remarkably close to zero. This aligns with the expectation of short-term financial market returns clustering around a zero mean; however, significant differences in the standard deviation reveal important distinctions in volatility among the three ETFs. Among the three, 0050 exhibits the highest standard deviation, in both weekly and daily returns. This finding indicates that, as a representative of the broad large-cap market, its price volatility is generally more pronounced. In contrast, 0056 and 00878 show lower volatility, reflecting the defensive and relatively stable nature of high-dividend strategies, making their prices less susceptible to sharp market fluctuations.

**Table 1: Summary Table of Descriptive Statistical Results**

Underlying Asset	Metric	Futures (F)	Spot (S)	NAV (Net Asset Value)
<b>0050 (Taiwan 50)</b>	Daily Return Mean	0.001446	0.001447	0.001434
	Daily Return Standard Deviation	0.012058	0.012233	0.012586
	Weekly Return Mean	0.006320	0.006361	0.006417
	Weekly Return Standard Deviation	0.026974	0.027020	0.028326
	5-Minute Return Standard Deviation	0.001547	0.001539	0.001717
<b>0056 (High Dividend)</b>	Daily Return Mean	0.000373	0.000376	0.000355
	Daily Return Standard Deviation	0.009561	0.009341	0.010080
	Weekly Return Mean	0.001807	0.001861	0.001775
	Weekly Return Standard Deviation	0.021887	0.021814	0.022225
	5-Minute Return Standard Deviation	0.001357	0.001202	0.001461
<b>00878 (Sustainable High Dividend)</b>	Daily Return Mean	0.000271	0.000271	0.000240
	Daily Return Standard Deviation	0.009025	0.008738	0.009737
	Weekly Return Mean	0.001470	0.001468	0.001421
	Weekly Return Standard Deviation	0.016415	0.016219	0.017756
	5-Minute Return Standard Deviation	0.001300	0.001219	0.001428

Furthermore, a comparison across price categories reveals that the volatility of NAV returns is generally slightly lower than that of spot and futures prices. This suggests that, short-run, market microstructure factors, such as noise from immediate exchange trading, changes in liquidity, and bid-ask spreads, may cause spot and futures prices to deviate from their intrinsic value. On the other hand, the standard deviation for 5-minute returns is the lowest across all ETFs, which is consistent with the known characteristic of high-frequency data, i.e. lower variance. Collectively, the table reveals a clear differentiation in risk structures: the broad market ETF (0050) carries higher risk, while the high-dividend ETFs (0056, 00878) show comparatively low volatility. The volatility of spot and futures prices typically exceeds NAV, reflecting the impact of trading dynamics on short-term prices.

This study sequentially performed an Augmented Dickey–Fuller (ADF) test on each data series (futures, spot, NAV) for the three selected ETFs, to examine their stationarity. The empirical results strongly and consistently indicate that all ETF series are stationary. Across all three data frequencies (weekly, daily, 5-minute), the t-statistics from the ADF tests are highly significant, and lower than the critical values for all significance levels, where the majority of p-values are  $< 0.001$ . The ADF test results provide sufficient evidence to reject the null hypothesis of a unit root, indicating the time series is stationary.

#### 4.2 VAR Model Results and Dynamic Interaction

The Vector Autoregression (VAR) empirical results for the three ETFs reveal that data frequency significantly determines the strength of market connectivity and predictive capability.

Table 2 presents the VAR results for the Yuanta Taiwan 50 ETF (0050). At a daily frequency, most market interactions appear relatively subdued. Most coefficients are statistically insignificant, except for a significant negative autocorrelation in the second-lagged futures returns. Most cross-variable impacts, such as futures on spot, or spot on NAV, fail to reach statistical significance. This evidence suggests that price discovery for 0050 is highly efficient and arbitrage opportunities are low, as cross-market information transmission is completed almost instantaneously intra-day, rather than persisting across days.

**Table 2: VAR result of 0050 ETF**

Variable	DAILY			5 MINUTES		
	Futures Return	Spot Return	NAV Return	Futures Return	Spot Return	NAV Return
<b>Futures Return (-1)</b>	-0.3495 (-0.71)	-0.2374 (-0.47)	0.5519 (1.54)	-0.0970*** (-3.69)	0.0350 (1.35)	0.0383** (2.19)
<b>Futures Return (-2)</b>	-0.9769** (-1.98)	0.2917 (0.57)	0.6138* (1.71)	-0.0860*** (-3.27)	0.062** (2.41)	0.0240 (1.37)
<b>Spot Return (-1)</b>	0.2138 (0.43)	-0.7913 (-1.54)	0.5466 (1.52)	0.3363*** (12.95)	-0.3894*** (-15.26)	0.0341** (1.97)
<b>Spot Return (-2)</b>	-0.7378 (-1.48)	0.1151 (0.22)	0.5562 (1.53)	0.1175*** (4.50)	-0.1169*** (-4.60)	0.0114 (0.66)
<b>NAV Return (-1)</b>	0.1193 (0.23)	-0.1605 (-0.30)	0.0138 (0.04)	0.4126*** (41.47)	0.0653** (2.33)	-0.4762*** (-25.03)
<b>NAV Return (-2)</b>	-0.8231 (-1.59)	0.4106 (0.77)	0.3257 (0.87)	0.1510*** (5.26)	0.0519* (1.86)	-0.2058*** (-10.84)

Notes:

1. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% levels, respectively.
2. The values in parentheses are the t-statistics.

In the 5-minute high-frequency data, there is a phenomenon that contrasts with conventional wisdom, where futures prices typically lead spot. Specifically, the first-lagged return of the 0050 spot price exhibits a highly significant and positive impact on current future returns. Furthermore, the influence of the previous period NAV return on futures price is even more substantial, indicating that futures price closely follows NAV direction in the ultra-short term. Conversely, the impact of lagged futures return on spot return is statistically insignificant, demonstrating only a marginal positive effect on NAV.

Table 3 exhibits the VAR results of 0056, which displays a slightly different finding to Table 2. In the 5-minute high-frequency data, 0056 futures returns (Lag 1), exhibit a highly significant and positive impact on both current spot return and NAV return. This indicates that market information is first reflected in the futures market, which subsequently drives spot price and NAV. Additionally, lagged spot significantly influences NAV, suggesting that spot price changes are incorporated into NAV. While lagged NAV returns positively affect spot returns, they show no significant impact on futures. These results demonstrate that 0056 follows a ‘futures-lead’ logic, where the futures market serves as the frontline for information transmission. This phenomenon is likely attributable to the composition of 0056 futures traders, or hedging demands, allowing the futures market to react to market shifts earlier than the spot market. However, in the daily data, unlike the insignificant performance of the 0050 ETF, 0056 exhibits a significant degree of co-movement.

**Table 3: VAR result of 0056 ETF**

Variable	DAILY			5 MINUTES		
	Futures Return	Spot Return	NAV Return	Futures Return	Spot Return	NAV Return
<b>Futures Return (-1)</b>	-1.1117*** (-3.29)	1.0134** (2.36)	0.2025 (0.70)	-0.1873*** (-17.91)	0.1587*** (8.79)	0.1459*** (9.58)
<b>Futures Return (-2)</b>	-0.6290* (-1.86)	0.3550 (0.85)	0.2451 (0.86)	-0.0840*** (-8.23)	0.0869*** (4.84)	0.0588*** (3.85)
<b>Spot Return (-1)</b>	-0.5297 (-1.59)	0.5002 (1.18)	0.1628 (0.57)	-0.0460*** (-4.79)	-0.1477*** (-9.00)	0.1660*** (11.96)
<b>Spot Return (-2)</b>	-0.3178 (-0.95)	0.0429 (0.11)	0.2261 (0.80)	-0.0110 (-1.16)	-0.068*** (-4.14)	0.0471*** (3.39)
<b>NAV Return (-1)</b>	-0.6396* (-1.76)	0.9968** (2.16)	-0.2189 (-0.70)	-0.0092 (-0.79)	0.0979*** (4.89)	-0.0901*** (-5.34)
<b>NAV Return (-2)</b>	-0.3526 (-0.97)	0.2564 (0.57)	0.0025 (0.01)	0.0091 (0.80)	0.0090 (0.45)	-0.0610*** (-3.57)

Notes:

1. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% levels, respectively.
2. The values in parentheses are the t-statistics.

Specifically, the prior day futures exert a significant positive driving effect on current spot returns. Furthermore, within the daily data, lagged NAV returns show a significant and positive impact on spot returns.

Table 4 shows the VAR estimation results for 00878. Regarding daily frequency, most coefficients are statistically insignificant, with notable exceptions being the first-lagged term of the futures itself (-0.7317), and also the impact of futures on NAV. In contrast, the 5-minute frequency yields coefficients that are nearly all significant at the 1% level. This further proves that the dynamics of price discovery and arbitrage are captured at much smaller time scales, rather than on a daily basis. Beyond the previous findings, 00878 shows a much stronger lead-lag relationship between NAV and spot price than 0050. Prior period NAV returns have a highly significant effect on current spot, indicating the presence of a powerful feedback loop.

This stronger correlation implies that 00878 investors are more sensitive to NAV transparency, and likely use it as a primary reference for their trading decisions.

**Table 4: VAR result of 00878 ETF**

Variable	DAILY			5 MINUTES		
	Futures Return	Spot Return	NAV Return	Futures Return	Spot Return	NAV Return
<b>Futures Return (-1)</b>	-0.7317** (-2.08)	0.2631 (0.61)	0.4334* (1.68)	-0.2689*** (-22.53)	0.2075*** (11.6)	0.1706*** (11.08)
<b>Futures Return (-2)</b>	-0.5980* (-1.72)	0.2612 (0.61)	0.2366 (0.9)	-0.1371*** (-11.89)	0.1274*** (7.11)	0.0513*** (3.31)
<b>Spot Return (-1)</b>	-0.2142 (-0.63)	-0.1508 (-0.36)	0.3479 (1.38)	-0.0080 (-0.69)	-0.2071*** (-11.96)	0.1843*** (12.35)
<b>Spot Return (-2)</b>	-0.2121 (-0.63)	-0.1069 (-0.26)	0.2209 (0.87)	-0.0082 (-0.73)	-0.551*** (-3.18)	0.0389** (2.59)
<b>NAV Return (-1)</b>	-0.2963 (-0.78)	-0.4201 (-0.89)	-0.1326 (-0.47)	0.0076 (0.57)	0.1805*** (8.88)	-0.1754*** (-10.02)
<b>NAV Return (-2)</b>	-0.3130 (-0.83)	0.1789 (0.39)	0.0144 (0.05)	0.0129 (0.98)	0.0860*** (4.21)	-0.1110*** (-6.33)

Notes:

1. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% levels, respectively.
2. The values in parentheses are the t-statistics.

Regarding the transmission of price information, for the futures, higher-frequency data is the decisive information leader; such that, as the price changes, it positively and significantly leads spot price and NAV. Concurrently, the spot market also acts as an information intermediary, imparting information to NAV. This precise transmission path is particularly pronounced in 0050 and 00878. Notably, 0056 exhibits a unique two-way information feedback loop in the high frequency data series, where spot and NAV also significantly influence futures. This suggests that the market integration and interaction is tightly coupled for large-cap weighted stocks, especially in the very short-run.

### 4.3 Granger Causality Test

Building upon the previous VAR results, the next phase of this research involves conducting a Granger Causality Test. Doing this clarifies the directional flow of information between futures, spot, and NAV markets for 0050, 0056, and 00878; thereby identifying which market serves as the true source of price discovery.

Table 5 shows the Granger causality test for 0050. At a daily frequency, no significant Granger causality exists among futures, spot, or NAV. This aligns with weak-form market efficiency theory, suggesting that information is fully assimilated by the market overnight, leaving no inter-day predictive power. However, examining the 5-minute high-frequency data, the information flow between markets becomes highly significant, revealing strong dynamic interactions. Specifically, the spot market exhibits a highly significant causal relationship with futures, indicating that movements in 0050's underlying components are the primary source of futures price volatility. The causal link from NAV to futures is the most prominent, suggesting that futures prices adjust by closely tracking NAV changes in the ultra-short term.

**Table 5: Granger Causality Test for 0050 ETF**

ETF / Frequency	Test Direction (Explanatory Variable → Dependent Variable)	Chi-square Statistic ( $\chi^2$ )	Degrees of Freedom (df)	p-value
<b>Daily Data</b>	Futures → Spot	1.032	2	0.597
	Futures → NAV	3.748	2	0.19
	Spot → Futures	3.585	2	0.167
	Spot → NAV	3.244	2	0.198
	NAV → Futures	3.582	2	0.167
	NAV → Spot	1.164	2	0.559
<b>5-Minute Frequency</b>	Futures → Spot	6.058	2	0.048**
	Futures → NAV	5.025	2	0.081*
	Spot → Futures	167.9	2	≤0.001***
	Spot → NAV	3.93	2	0.14
	NAV → Futures	209.4	2	≤0.001***
	NAV → Spot	6.621	2	0.036**

Notes:

1. \*\*\*, \*\*, and \* denote the rejection of the null hypothesis (i.e., Granger causality exists) at the 1%, 5%, and 10% significance levels, respectively.
2. ≤0.001 indicates that the p-value is less than or equal to 0.001.
3. Null Hypothesis (H<sub>0</sub>): The explanatory variable does not Granger cause the dependent variable. If the p-value is significant, H<sub>0</sub> is rejected, implying the existence of a causal relationship.

While futures also lead spot and NAV, the statistical significance and magnitude of this relationship is far weaker than the influence exerted by the latter on the former. As indicated in Table 5, the price discovery function of 0050 is not dominated by futures; instead, it is driven by the spot and NAV markets, with futures acting as a highly sensitive follower that continuously corrects its pricing.

With respect to the Granger causality test for 0056, Table 6 reveals that at the daily frequency, only the ‘futures-leads-spot’ and ‘NAV-leads-spot’ relationships exhibit marginal significance, and this is only at the 10% level. This suggests that from a long-term perspective, the causal linkages among 0056’s markets are not pronounced, which is a similar finding to 0050. However, at the 5-minute frequency, nearly all causal paths reach extreme significance at the 1% level, indicating that 0056’s price information transmission primarily occurs during ultra-short-term intra-day trading. Specifically, futures exert a powerful causal influence on both spot and NAV, supporting the role of futures as a primary leading indicator. Additionally, the spot market triggers movements in both futures and NAV. Notably, the impact of the spot market on NAV yields the highest coefficient in Table 6, showing that fluctuations in 0056’s spot price significantly dictate NAV expectations.

**Table 6: Granger Causality Test for 0056 ETF**

ETF / Frequency	Test Direction (Explanatory Variable → Dependent Variable)	Chi-square Statistic ( $\chi^2$ )	Degrees of Freedom (df)	p-value
<b>Daily Data</b>	Futures → Spot	5.587	2	0.061*
	Futures → NAV	0.931	2	0.628
	Spot → Futures	2.608	2	0.271
	Spot → NAV	0.751	2	0.687
	NAV → Futures	3.149	2	0.207
	NAV → Spot	4.573	2	0.093*
<b>5-Minute Frequency</b>	Futures → Spot	87.2	2	$\leq 0.001^{***}$
	Futures → NAV	94.4	2	$\leq 0.001^{***}$
	Spot → Futures	23.17	2	$\leq 0.001^{***}$
	Spot → NAV	143.3	2	$\leq 0.001^{***}$
	NAV → Futures	1.472	2	0.479
	NAV → Spot	24.2	2	$\leq 0.001^{***}$

Notes:

1. \*\*\*, \*\*, and \* denote the rejection of the null hypothesis (i.e., Granger causality exists) at the 1%, 5%, and 10% significance levels, respectively.
2.  $\leq 0.001$  indicates that the p-value is less than or equal to 0.001.
3. Null Hypothesis (H0): The explanatory variable does not Granger cause the dependent variable. If the p-value is significant, H0 is rejected, implying the existence of a causal relationship.

Table 7 relates to ETF 00878, and at a daily time interval, reveals no significant causal relationships among the futures, spot, or NAV returns. This indicates that the market for 00878 is highly efficient, as all relevant information is fully digested and reflected in the price by the end of each trading day, eliminating inter-day lead-lag relationships. However, for the 5-minute interval, 00878 exhibits an exceptionally strong information transmission path, compared to the other two ETFs. Specifically, the 00878 futures market acts as the sole price discoverer. New market information is first reflected in the futures price, which then unilaterally drives both spot and NAV, with the latter two failing to exert any reverse influence. Furthermore, a clear bi-directional feedback mechanism is observed between the spot and NAV markets.

**Table 7: Granger Causality Test for 00878 ETF**

ETF / Frequency	Test Direction (Explanatory Variable → Dependent Variable)	Chi-square Statistic ( $\chi^2$ )	Degrees of Freedom (df)	p-value
<b>Daily Data</b>	Futures → Spot	0.561	2	0.755
	Futures → NAV	2.859	2	0.239
	Spot → Futures	0.572	2	0.751
	Spot → NAV	2.007	2	0.367
	NAV → Futures	0.941	2	0.625
	NAV → Spot	0.807	2	0.668
<b>5-Minute Frequency</b>	Futures → Spot	148.17	2	≤0.001***
	Futures → NAV	122.72	2	≤0.001***
	Spot → Futures	0.828	2	0.661
	Spot → NAV	154.62	2	≤0.001***
	NAV → Futures	1.097	2	0.578
	NAV → Spot	81.25	2	≤0.001***

Notes:

1. \*\*\*, \*\*, and \* denote the rejection of the null hypothesis (i.e., Granger causality exists) at the 1%, 5%, and 10% significance levels, respectively.
2. ≤0.001 indicates that the p-value is less than or equal to 0.001.
3. Null Hypothesis (H<sub>0</sub>): The explanatory variable does not Granger cause the dependent variable. If the p-value is significant, H<sub>0</sub> is rejected, implying the existence of a causal relationship.

#### 4.4 Discussion of Empirical Results

The combined findings from the VAR and Granger Causality tests provide a comprehensive view of ETF market dynamics. The results consistently show that futures, spot, and NAV series for all three ETFs are stationary, confirming the existence of a stable short-term equilibrium structure in the market.

Crucially, price information from the futures market maintains a critical guiding role in the price discovery process, which is particularly visible in the high-frequency data series. The futures market instantaneously reflects new market information, and exhibits a significant leading effect on both spot price and NAV. This clearly demonstrates the strong price discovery function of the futures market over the short term. Conversely, the adjustment speed of spot and NAV prices is comparatively slow, indicating that there is a limit in the immediacy of their reaction to new information. This lag may be attributable to factors such as the time delay in fund NAV calculation and other market transaction costs. These findings align with conclusions already found in the literature, such as Tse (1995), Alphonse (2000), and Kavussanos et al. (2008). This study finds that the futures market leads information transmission and price formation. It is noteworthy, however, that this futures-led pattern does not apply uniformly across all ETFs. For 0050, the empirical results indicate a reverse dynamic, where spot prices and NAV more strongly Granger-cause futures prices. This may be attributed to the high liquidity and transparency of 0050's underlying large-cap constituents, where spot market information is rapidly incorporated and subsequently reflected in futures pricing.

The empirical results of this study not only reveal the structural characteristics of the ETF market, but also offer strategic implications: For investors, futures prices serve as a leading indicator of short-term market changes, aiding in the prediction of spot and NAV trends. This information is directly applicable to short-term trading and arbitrage decisions. For regulators, the immediacy of the price signals from the futures

market offers a vital reference for assessing market efficiency and stability, thereby strengthening the transparency and efficiency of operations in complex financial markets.

## 5. Conclusion

This study examined Taiwanese ETFs and investigated the sequence of price discovery transmission throughout the market, with respect to futures prices, spot prices, and Net Asset Value (NAV). Taking the Yuanta Taiwan 50 (0050), Yuanta High Dividend (0056), and Cathay Sustainable High Dividend (00878) ETFs as a representative study cohort, Vector Autoregression (VAR), and Granger Causality tests were used to probe for possible lead-lag relationships using differing time intervals.

The empirical results revealed a significant contrast between differing observation frequencies. At a daily frequency, all three ETFs exhibited weak-form market efficiency, where inter-market lead-lag relationships vanish as information is fully assimilated within the trading day. However, at 5-minute data intervals, the study captured the dynamic interactions and hidden information flows that are otherwise masked by daily observations. This justifies the need for further high-frequency analysis in market microstructure research. Overall, the price discovery patterns differ across the three ETFs. For 0056 and 00878, the futures market plays a clear leading role in price discovery, with futures prices significantly driving spot prices and NAV, particularly at the 5-minute frequency. However, for 0050, the relationship is more complex: spot prices and NAV exert a stronger and more significant causal influence on futures prices, suggesting that this broad market ETF's price discovery is driven primarily by the underlying spot market rather than the futures market. This heterogeneity underscores that the dominance of the futures market in price discovery is not universal, but contingent upon the specific characteristics of each ETF.

In contrast, spot and NAV returns largely exhibit a passive reaction, lagging changes in futures prices. A comparison of the different data frequencies reveals that the Granger Causality test results for daily data are mostly insignificant, suggesting that price interaction between the markets is indistinct at lower frequencies. However, in the 5-minute data frequency, the leading effect of futures on spot and NAV is robust. This phenomenon shows that information transmission is time-sensitive, and higher frequency data is essential when trying to capture the direction of information flow. In summation, this study confirms that ETF futures inform price and NAV, and also that higher frequency data possesses superior explanatory power for the direction of information flows.

Based on the findings in this study, it is possible to make several recommendations. First, for investors, futures prices should be utilized as a leading indicator of short-term market movements. Investors can leverage this information to develop short-term trading or arbitrage strategies, particularly in high-frequency trading environments. Second, for regulators, continuous monitoring of price transmission between ETF futures and spot markets can be used to maintain market stability during periods of high volatility. If futures consistently lead, and diverge significantly from spot prices, it may signal a market structure imbalance, or some amplification in speculative trading. Such situations may be mitigated by liquidity management or improved information disclosure rules.

Finally, this research suggests several directions for future study. Researchers could extend the investigation to different types of ETF (e.g., thematic, sustainable, or leveraged ETFs) or incorporate cross-market comparisons using multi-market data (e.g., Taiwanese and international ETF markets). Furthermore, future work could adopt even higher frequency data, utilizing advanced models such as the Dynamic Conditional Correlation GARCH (DCC-GARCH) or Time-Varying Granger Causality, to deeply analyze the immediacy and stability of information transmission. Such efforts would further enhance our understanding of ETF market structures, liquidity, and the efficiency of price discovery. Future research could also extend the sample period, broaden the ETF sample range, or combine multivariate volatility models with high-frequency data.

## References

- Alphonse, P. (2000). Efficient price discovery in stock index cash and futures markets. *Annales d'Economie et de Statistique*, (57/58), 177–188.
- Atilgan, Y., Demirtas, K. O., Gunaydin, A. D., & Oztekin, M. (2022). Price discovery in emerging market ETFs. *Applied Economics*, 54(47), 5476–5496.
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a), 427–431.
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 37(3), 424–438.
- Hao, J., Xiong, X., He, F., & Ma, F. (2019). Price discovery in the Chinese stock index futures market. *Emerging Markets Finance and Trade*, 55(13), 2982–2996.
- Ho, W.J., & Tseng, C.-W. (2007). Price discovery of the Taiwan 50 Index ETF. *Journal of Chinese Economic Research*, 5(1), 87–107. (in Chinese).
- Hsieh, W.-L. (2002). Price discovery, information transmission, and market integration: Evidence from Taiwan's stock index futures market. *Journal of Financial Studies*, 10(3), 1–31. (in Chinese).
- Hu, Y., Hou, Y. G., & Oxley, L. (2020). What role do futures markets play in Bitcoin pricing? Causality, cointegration and price discovery from a time-varying perspective? *International Review of Financial Analysis*, 72, Article 101569.
- Kaur, P., & Singh, J. (2022). Investigating the informational efficiency of exchange traded funds (ETFs): Empirical evidence from Indian equity ETFs. *Finance India*, 36(1), 193–208.
- Kavussanos, M. G., Visvikis, I. D., & Alexakis, P. D. (2008). The lead–lag relationship between cash and stock index futures in a new market. *European Financial Management*, 14(5), 1007–1025.
- Kia, K., Song, V. Y., & Xu, K. (2024). Price discovery in Bitcoin ETF market. Available at SSRN 4776832.
- Ji, H., Naeem, M., Zhang, J., & Tiwari, A. K. (2024). Dynamic dependence and spillover among the energy related ETFs: From the hedging effectiveness perspective. *Energy Economics*, 136, 107681.
- Poutachidou, N., & Koullis, A. (2025). The investment styles and performance of AI-related ETFs: Analyzing the impact of active management. *FinTech*, 4(2), 20.
- Securities Investment Trust & Consulting Association (SITCA). (2025). Opportunities and challenges of Taiwan ETF market in 2025 [Presentation slides].
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 48(1), 1–48.
- Tse, Y. K. (1995). Lead–lag relationship between spot index and futures price of the Nikkei Stock Average. *Journal of Forecasting*, 14(7), 553–563.