

One Instrument, Two Technologies: Price Discovery Between Tokenized and Traditional Equity Markets

Kaitao Lin*

August 2026

*Senior Financial Economist, World Federation of Exchanges. Email: klin@world-exchanges.org. The views expressed in this paper are my own and do not reflect those of the World Federation of Exchanges or its member institutions. Any errors that remain are my sole responsibility.

One Instrument, Two Technologies: Price Discovery Between Tokenized and Traditional Equity Markets

Abstract

Tokenized equities place a claim on a listed stock in a around-the-clock trading, on-chain market connected to the underlying by a redeemable custodial claim. Using a daily panel of 61 token–underlying pairs over the market’s first eleven months, we find that the no-arbitrage condition holds on a daily frequency, as prices are cointegrated for more than 80% of pairs, with a long-run relationship almost exactly one-for-one in log prices. We also find that the pricing tether is anchored to the traditional exchange market. The token does the majority of the adjustment to disequilibrium, and permanent–transitory and information-share decompositions attribute the majority of price discovery to the underlying equity. Cointegration is correlated by the depth of the equity market, and cross-pair variation in price discovery. Overall, tokenized equities function as a distribution channel for prices formed in the regulated market.

Keywords: tokenization, price discovery, cointegration, information shares, no-arbitrage

JEL classification: G12, G14, G15, G18, G23

1 Introduction

In July 2025, a set of publicly traded U.S. equities began trading around the clock on cryptocurrency platforms in the form of tokenized equities. The instruments, known as xStocks, are blockchain tokens issued against custodied shares of the stocks and ETFs they reference. Each token is collateralized one-for-one by the underlying security, can be held in a private digital wallet, and trades around-the-clock, on venues outside the traditional equity market structure. The claim on the same expected return and cash flow is thereby priced in two very distinct markets—a regulated exchange operating under disclosure and surveillance rules, and an on-chain market open at all hours, an order of magnitude thinner, and connected to the first only by a redeemable custodial claim.

This arrangement leads to two research questions we intend to address in this paper. The first one is whether the no-arbitrage condition holds across the token-underlying pair, as two claims to the same value should command the same price. The evidence from cryptocurrency markets shows the same digital asset routinely trades at materially different prices across venues (Makarov and Schoar, 2020), and even tokens engineered to hold a fixed value drift from it (Griffin and Shams, 2020; Lyons and Viswanath-Natraj, 2023). Second, if the two prices share a common long-run value, which market sets it? Does the common efficient price form in the regulated equity market, with the token following, or does the continuously trading token incorporate information first?

In this paper, we answer both questions using a daily panel of 61 matched token-underlying pairs from July 2025 through May 2026. The tokenized equities¹ are issued by Backed and traded on crypto trading platforms, including Kraken and Bybit, among others, and the underlying equities are the corresponding equities and ETFs listed and traded on U.S. stock exchanges. Regarding the no-arbitrage condition, our result shows that the tether holds. Log token and underlying prices are individually integrated of order one,

¹We refer to such digital tokenized claims as tokenized equities. The U.S. SEC classifies them as Synthetic Tokenized Securities. See <https://www.sec.gov/newsroom/speeches-statements/corp-fin-statement-tokenized-securities-012826-statement-tokenized-securities>

and cointegration tests show a common trend for about 80% of the sample pairs under both the residual-based test of Engle and Granger (1987) and a deviation unit-root test (Taylor and Taylor, 2004). The panel Fisher test (Maddala and Wu, 1999) confirms that the two markets are cointegrated when tested jointly across all sample pairs. Moreover, the estimated long-run relationship is almost exactly one-for-one in log prices, as the backing predicts. In the cross-sectional analysis, the token-underlying pairs with higher trading volume in the underlying equity and lower trading volume in the tokenized market tend to be cointegrated.

Regarding price discovery, we find that the tether is anchored in the traditional market. The evidence comes from three related but conceptually distinct measures, each investigating a different aspect of the way common price forms. The first one is the vector error correction model (VECM), which estimates each market's reaction when deviation takes place between the two prices. Because the two prices are cointegrated, any deviation between them is temporary, and the adjustment coefficients in the VECM measure the division of the burden of eliminating that price deviation. On the one hand, a market that moves to close the gap suggests, in effect, that its price is inefficient. On the other hand, a market that does little adjustment, while the other converges toward it, is the anchor of the common price. In our estimates, the token does the majority of this adjustment. Its adjustment coefficient averages -0.602 and is statistically significant in about 73% of the sample pairs. When the token price closes above its underlying price, more than half of the price discrepancy is eliminated by the token's own price falling back within the next trading day. The underlying, by contrast, responds less to the disequilibrium. The underlying equity's adjustment coefficient averages 0.228 and is only statistically significant in 27% of the sample pairs. The underlying equity behaves as the anchor around which the token oscillates.

The second measure, the permanent–transitory decomposition of Gonzalo and Granger (1995), combines the VECM estimated adjustment speeds into a measure of the proportion

of common price each market is attributed to. The decomposition intuition follows that the common permanent component is built from the two markets in inverse proportion to how much each one adjusts. A market that never moves when a gap opens contributes everything to the common trend, because its price is the anchor. A market that does all of the correcting contributes nothing, because its movements are precisely the transitory noise being corrected. Our result shows that the decomposition attributes more than 80% of the common permanent component to the equity for the median pair. The token is, in other words, a satellite, as its price is informative mainly as a noisy reflection of the value formed in the traditional equity market.

The third measure, the information share of Hasbrouck (1995), estimates whose price movement forms the common price. Rather than apportioning the correction of price deviation, the information share decomposes the variance of innovations to the common efficient price into the contribution of each market's own shocks. A market with a large information share is one whose innovation tend to become permanent in the common price. Because the two markets' daily innovations are contemporaneously correlated, the decomposition is not unique, and we report bounds rather than point estimates. The results nonetheless show that the underlying equity leads in about 20% of the sample pairs by the most conservative criterion with its lower bound alone exceeding 0.5. Yet, in not a single pair does the token lead, as the equity's upper bound never falls below 0.5. The information shares thus corroborate the direction of the adjustment-based evidence that the underlying equity market contribute more to the common price of token-underlying pairs.

The result of the underlying equity leading price discovery is conservative by construction. The token's closing price is stamped at midnight Coordinated Universal Time (UTC) several hours after the U.S. equity close at 4:30 p.m. Eastern Time (ET). Also, the token trades around-the-clock in between. Therefore, between any two daily observations the token has a larger window to incorporate new information. The features of the data tilt

the measurement toward attributing price discovery to the token, and the equity leads regardless strengthens the conclusion.

Further cross-sectional evidence completes the picture. Whether a pair is cointegrated is associated with the depth of the underlying equity market. We also find that the contribution of price discovery across pairs is correlated with the relative volatility of the two legs. The underlying equity contributes more to the price discovery for the token-underlying pairs that have lower volatility in the underlying equity market and higher volatility in the tokenized market. This finding suggest that excess and correctable movement on a leg lowers that leg's contribution to price discovery. Both findings point to the same direction. The token market functions as a distribution channel for exposures whose prices are formed in the traditional exchange market.

Our findings contribute to three strands of literature. First, we extend the price discovery literature on multi-market claims—cross-listed stocks, depositary receipts, and index products—in which discovery concentrates in the more liquid, established venue (Garbade and Silber, 1983; Hasbrouck, 1995; Eun and Sabherwal, 2003; Hasbrouck, 2003). Our contribution is its setting. The following venue is an around-the-clock on-chain market connected to the leader only by a redeemable custodial claim, and that claim suffices to keep essentially all discovery in the underlying. Second, we extend the ETF and closed-end fund literature on backed claims, in which creation–redemption arbitrage tethers the traded price to the underlying value but slackens where arbitrage is costly (Ackert and Tian, 2000; Engle and Sarkar, 2006; Petajisto, 2017; Ben-David, Franzoni, and Moussawi, 2018). Tokenized equities behave less like a new asset class than like a familiar backed claim operating through a settlement technology. Third, we contrast with the cryptocurrency literature documenting fragmented, leaderless pricing across venues (Makarov and Schoar, 2020; Brandvold et al., 2015). The one-for-one backing with a redemption channel could impose on an on-chain market the pricing discipline that native crypto assets lack.

The paper closest to ours is Cong et al. (2025), an early and wide-ranging assessment of tokenized stocks documenting that token prices track their underlyings during exchange hours and that off-hour token returns predict the underlying’s subsequent opening return. Their evidence establishes the intertemporal dimension of price discovery, documenting that tokens extend the incorporation of information across periods when the primary market is closed. Our paper addresses the complementary cross-market dimension. We investigate when both markets are observed at a common frequency, which one sets the common price by using the error correction model and information-share measures that Cong et al. (2025) identify as a natural next step. The two sets of findings are mutually reinforcing. Cong et al. (2025)’s off-hour pass-through is what a satellite market relaying information toward the equity’s next price would produce, and our results confirm that the equity anchors the common value to which both prices revert.

The result carries significant policy implications. For exchanges, it suggests that the value of a listing venue lies in the concentration of informed order flow, which does not migrate merely because an always-open alternative exists. Tokenization as currently constituted is not disintermediation but distribution, carrying the exchange’s prices to new hours and new investors while contributing less to the price discovery. For regulators, the same fact cuts both ways. The informational integrity of token prices rests on a market inside the existing regulatory perimeter. However for this very reason, tokenized equity investors receive the regulated market’s prices without its protections and oversights.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature. Section 3 describes the institutional design of tokenized equities. Section 4 describes the data. Section 5 tests the no-arbitrage condition through cointegration. Section 6 estimates the division of price discovery. Section 7 examines its cross-pair determinants, and Section 8 concludes.

2 Related Literature

Our central question—where price is discovered when a single fundamental value trades in more than one market—has a long lineage in empirical finance. Studies of cross-listed stocks and American Depositary Receipts (ADRs) consistently find that price discovery gravitates toward the more liquid and established venue, typically the home market, even when a second market trades the same claim during overlapping hours (Garbade and Silber, 1983; Werner and Kleidon, 1996; Eun and Sabherwal, 2003; Grammig, Melvin, and Schlag, 2005; Menkveld, 2008), and the same finding holds for a basket or derivative instruments and their underlyings (Hasbrouck, 2003; Chakravarty, Gulen, and Mayhew, 2004). Our finding that the underlying U.S. equity, rather than the token, is the primary site of price discovery places tokenized equities squarely within this setting that the deeper and longer-established market leads.

A token backed by custodied shares is, economically, close to an exchange-traded fund (ETF) share backed by a basket. In both cases, a creation-and-redemption mechanism is meant to tether the traded price to the value of the underlying. The ETF and closed-end-fund literature shows that this tether is tight for liquid, easily arbitrated instruments but slackens as arbitrage becomes more costly, producing persistent premiums and discounts for the least liquid funds (Ackert and Tian, 2000; Engle and Sarkar, 2006; Petajisto, 2017; Ben-David, Franzoni, and Moussawi, 2018). Our results extend this literature to on-chain claims. No-arbitrage condition holds for the majority of pairs, and the tokens that decouple most from their underlyings are concentrated among the tokens with lower trading volume. In this sense, tokenized equities behave less like a new asset class than like a familiar backed claim operating through an unfamiliar settlement technology.

That the tether holds is not a foregone conclusion, because the evidence from cryptocurrency markets suggests it could fail. In these markets the same asset trades on multiple venues, and prices diverge frequently and sizably, because moving capital across venues is slow and costly (Makarov and Schoar, 2020; Lin, 2026a). Also tokens engi-

neered to hold a fixed value are documented to drift from it (Griffin and Shams, 2020; Lyons and Viswanath-Natraj, 2023). Tokenized equities carry the arbitrage tether of a backed claim but trade in the frictional, around-the-clock environment of cryptocurrency. Our finding suggests that the tether largely holds and the traditional equity market retains price-discovery primacy.

Closest to our work, Cong et al. (2025) provide an early and wide-ranging empirical assessment of tokenized stocks across the two dominant issuers, documenting that token prices track their underlyings closely during exchange hours, respond to earnings news released outside U.S. trading hours, and that off-hour token returns predict the underlying’s subsequent opening return. Their evidence establishes the *intertemporal* dimension of price discovery in this market, as tokens extend the incorporation of information across periods in which the primary market is closed. Our paper addresses the complementary *cross-market* dimension. When both markets are open and observed at a common frequency, which market sets the common price? We accordingly estimate pair-level vector error correction models and decompose the common price using the Gonzalo–Granger and Hasbrouck measures—an analysis that Cong et al. (2025) identify as a natural next step for this literature. The two papers’ findings are mutually reinforcing. Their evidence that off-hour token returns pass through nearly one-for-one into the next open is consistent with our finding that the token is a satellite of the equity rather than an independent site of price formation. The token relays information toward the equity’s next price, while the equity anchors the common value to which both prices revert.

3 Institutional background

In this paper, we investigate the xStocks, which are blockchain-based tokenized equities designed to track the price of publicly traded U.S. stocks or ETFs. The tokens are issued by Backed, and each token is collateralized one-for-one by the corresponding underlying

security held with a regulated custodian, with segregated custody accounts.² The product was launched in July 2025, which determines the start of our sample. Although each token is backed by the underlying share, holding an xStock does not confer ownership, voting, or other shareholder rights in the underlying company.

A few features distinguish the token from a conventional claim on the same security. First, the token is a self-custodial instrument. It can be held in a private wallet and transferred across platforms, and it exists natively on multiple blockchains, including Solana and Ethereum. Second, the tokenized equities trade 24/7.³ While the underlying trades only during the hours of its listed exchange, xStocks trade around the clock on cryptocurrency platforms, including weekends and holidays.⁴ Moreover, the one-for-one correspondence between token and the underlying equity adjusts correspondingly during corporate actions. Dividends, stock splits, and reverse splits are passed through to holders by adjusting the token balances so that a holder's position maintains one-for-one exposure to the underlying. Moreover, xStocks are only available to non-U.S. retail clients in select countries.

Tokens enter and exit circulation through a market operated by the issuer, in which participants can create new tokens against the underlying or redeem tokens at the prevailing market price of the underlying security. This mechanism is the economic tether between the two markets, and it operates similarly to the creation–redemption process that keeps an ETF near the value of its basket portfolio. If the token is overpriced relative to the underlying, a participant can create tokens at the underlying's price and sell them at the token's price, and vice versa. However, while the token itself trades 24/7, issuance and redemption operate only 24/5, aligned with U.S. business days. The arbitrage channel is therefore closed during the non-exchange hours in which only the token market is open and token prices can drift on news that the equity cannot yet reflect.

²Institutional details in this section are drawn from the xStocks documentation at <https://docs.xstocks.fi/docs> and the issuer's legal documentation at <https://assets.backed.fi>.

³According to Kraken, tokenized tokens are traded 24/5 on its platform and 24/7 on-chain.

⁴For a discussion on extending exchange trading hours, see Lin (2026b).

4 Data

To carry out the analysis, we gather a daily panel data covering 79 tokenized stocks and ETFs issued by Backed and traded on the cryptocurrency platform Kraken and their corresponding underlying U.S.-listed securities. The sample spans from July 1, 2025 to May 31, 2026. For the tokenized equities, we obtain the daily closing prices and trading volume from Yahoo Finance. Data on the corresponding underlying U.S.-listed stocks and ETFs—again, daily closing prices and volumes—are drawn from Capital IQ. To control for market-wide conditions, we supplement the trading data with data from the Federal Reserve Economic Data (FRED): the three-month Treasury bill rate, which proxies for the risk-free rate; the S&P 500 index, which captures the aggregate equity market; and the CBOE Volatility (VIX) Index, a measure of market volatility.

Each tokenized equity is matched one-to-one to its underlying stock or ETF on the basis of the reference security it is designed to track, yielding a set of token–underlying pairs. Because the empirical strategy relies on per-pair time-series estimation, we require each pair to have at least 60 daily observations over the sample window. Pairs with shorter histories, typically tokens listed late in the sample period, are excluded. This filter reduces the 79 pairs of token-underlying to a final sample of 61 matched pairs, forming a panel of 10,848 pair-days.⁵

The two legs of each pair are observed at different points in the day. The tokenized price is recorded just before midnight at 23:59:59 UTC, whereas the underlying equity price is the 16:00 ET close, which corresponds to roughly 20:00–21:00 UTC depending on daylight saving time. The token closing price is therefore stamped several hours after the equity close. Therefore, for any given calendar day, the token price has the advantage

⁵We screen each pair on the closing-price ratio $\rho_t = P_t^x / P_t$, which should stay near one under one-for-one backing, flagging any pair with $\max \rho_t / \min \rho_t > 1.5$, and we make three corrections. For NFLX and TQQQ the token series is not adjusted for a stock split (ten-for-one and two-for-one, respectively); we rescale the token price and volume to post-split terms, exclude a residual window of TQQQ days on which the source mixes scales, and end the NFLX series on May 8, 2026, when its price feed switches to a contaminated multi-venue aggregate. For NET, ρ_t averages roughly 0.003, indicating a mismatched series, and we drop the pair.

of reflecting more recent information than its underlying equity does. This timing wedge works against the paper’s central result. Because the token price has the opportunity to incorporate same-day information that the earlier equity close cannot, the mechanism is, if anything, tilted toward attributing price discovery to the token. A finding that the underlying equity leads price discovery is therefore conservative. The same feature also limits what we can identify. Because prices are sampled once per day, the analysis speaks to price discovery at the daily horizon, and intraday dynamics between the two markets are not identified in this paper.

Table 1 reports sample means, standard deviations, and skewness of the return, trading volume, and liquidity measure for the two legs of each pair. We also report the underlying-minus-token difference in means and a paired t-test of that difference. Average daily returns are similar across the two markets—0.043% for the underlying and 0.044% for the token, a difference of -0.001 percentage points that is statistically indistinguishable from zero. This result is consistent with the token and its underlying being claims on the same return or future cash flow. The two return distributions are also alike, each only mildly right skewed (1.147 for the underlying and 0.608 for the token), with the token being more volatile (a standard deviation of 3.511% against 3.061%).⁶

Table 1. Summary statistics: underlying versus tokenized securities

This table reports pooled summary statistics for the four variables listed, computed separately for the underlying U.S.-listed securities and their tokenized counterparts across all 10,848 pair-days in the sample (61 matched token–underlying pairs observed daily from July 2025 through May 2026). For each variable the table reports the mean of each leg, the mean of the underlying-minus-token difference, and the standard deviation and skewness of each leg. Return is the daily return in percent. Share Volume is daily trading volume in millions of shares (or units, for ETFs). Dollar Volume is daily dollar trading volume in billions of U.S. dollars. Amihud is the illiquidity measure of Amihud (2002), scaled by 10^6 . The Mean Difference column also reports the significance of paired t-test. Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Variable	N	Mean			Std. Dev.		Skewness	
		Underlying	Tokenized	Difference	Underlying	Tokenized	Underlying	Tokenized
Return (%)	10,848	0.043	0.044	-0.001	3.061	3.511	1.147	0.608
Share Volume (million)	10,848	25.424	2.884	22.540***	36.196	6.090	3.106	3.056
Dollar Volume (\$ billion)	10,848	5.994	0.815	5.179***	10.789	1.960	3.354	4.742
Amihud ($\times 10^6$)	10,848	0.005	9.478	-9.474	0.047	466.144	17.062	98.773

The two markets diverge sharply, however, in trading activity. Average daily share vol-

⁶The difference is statically significant at the 1% level in a F-test.

ume in the underlying market is roughly nine times that of its tokenized counterpart (25.4 against 2.9 million) and average daily dollar volume roughly seven times as large (\$5.99 against \$0.82 billion). Both differences are significant at the 1% level. The Amihud (2002) illiquidity ratio, the average absolute return per dollar of volume, is far larger for the tokens (9.478 against 0.005, $\times 10^6$). This gap is large in economic terms but statistically insignificant.

5 Cointegration and the No-Arbitrage Condition

The no-arbitrage condition is among the most important propositions in asset pricing. Two claims to the same future cash flows must command the same price, since any persistent deviation could generate a riskless profit. Here the two claims are a tokenized equity and the underlying stock or ETF it is designed to mirror. As the token is backed by the underlying share and engineered to track its value, the two should trade at the same price up to transaction costs, and any price deviation between them should be temporary. Testing whether this holds is important, as a tokenized equity is economically useful only insofar as it reproduces the exposure of the security it references, and a token that departs systematically from its underlying saddles its holders with unintended basis risk and fails to deliver the very exposure that it is designed to provide.

However, the no-arbitrage condition is not guaranteed by the token's design. Unlike a futures contract, which converges to the underlying at delivery, a token has no maturity that mechanically forces the convergence of the two prices. The only tie is arbitrage, which operates through a restricted creation-and-redemption mechanism and is subject to on-chain transaction costs and possible constraints on short-selling the token. These frictions can prevent arbitrageurs from closing the price deviation immediately, so arbitrage opportunities could open and persist. The evidence from cryptocurrency markets shows that the same digital asset routinely trades at materially different prices across

venues (Makarov and Schoar, 2020), specially when there is delay in blockchain settlement (Hautsch, Scheuch, and Voigt, 2024; Lin, 2026a), and even tokens engineered to hold a fixed value drift from it (Griffin and Shams, 2020; Lyons and Viswanath-Natraj, 2023).

We investigate the no-arbitrage condition through cointegration. We first carry out the augmented Dickey–Fuller (ADF) tests for each of the 61 filtered token-underlying pairs for unit roots. Table A1 in the Appendix reports the p -values from the tests of the null hypothesis of a unit root, applied to the daily log closing price of each security in levels and in first differences. For each matched pair, the test is run separately on the underlying U.S.-listed security and on its tokenized counterpart. The results show that both log prices, $\ln P_t$ for the underlying and $\ln P_t^x$ for the token, are integrated of order one (I(1)). Under the 5% significance level, the joint I(1) criterion—failure to reject a unit root in levels together with rejection in first differences—is satisfied by 60 of the 61 underlying series (98.4%) and 57 of the 61 token series (93.4%).

We then test for the cointegration of the two pricing series. If a linear combination of the two I(1) prices is stationary, the pair shares a single common stochastic trend and the deviation from it is transitory. Cointegration is therefore the empirical test of the long-run no-arbitrage condition. We test cointegration for the underlying and token prices with two complementary tests—the Engle and Granger (1987) test and the price deviation unit root test (Taylor and Taylor, 2004).

Following the two-step procedure of Engle and Granger (1987), for each underlying-token pair, we estimate the cointegrating regression $\ln P_t^x = c + \beta \ln P_t + u_t$ by Ordinary Least Squares (OLS) and apply an ADF test to the estimated residuals $\hat{u}_t$. Stationary residuals imply the prices are tied together in the long run, so rejecting the unit-root is evidence of cointegration. A summary of the test results is reported in the first row of Table 2. At the 5% significance level, the residual-based test rejects the null hypothesis of no coin-

tegration for 50 of the 61 pairs (82.0%).⁷ In other words, the majority of our sample underlying–token pairs are cointegrated.

In addition, following Taylor and Taylor (2004), we impose the one-for-one backed relationship between the token and the underlying by forming the log price deviation as

$$dev_t = \ln P_t - \ln P_t^x, \quad (1)$$

which fixes the cointegrating vector at $(1, -1)$. Then, we test dev_t for a unit root. Because the price deviation is specified rather than estimated, the test is invariant to normalization and recovers power that the two-step procedure loses in estimating the slope, as in the Engle–Granger test. A stationary dev_t suggests departures between the token and underlying prices are corrected, rather than accumulated. As summarized in Table 2, the unit-root null is rejected for 48 of the 61 pairs (78.7%) at the 5% level, closely tracking the Engle–Granger results.

Table 2. Cointegration tests

This table summarizes pair-level cointegration tests for the 61 matched token–underlying pairs. The first row applies the two-step procedure of Engle and Granger (1987): the cointegrating regression $\ln P_t^x = a + b \ln P_t + u_t$ is estimated by OLS and an ADF test is applied to the residuals. The second row imposes the cointegrating vector $(1, -1)$ and applies an ADF test directly to the log price deviation $dev_t = \ln P_t - \ln P_t^x$, following Taylor and Taylor (2004). Columns report the number of pairs rejecting the unit-root null at each significance level. The final two columns report the Maddala and Wu (1999) Fisher statistic, $-2 \sum_i \ln(p_i) \sim \chi^2(2N)$.

Test	N	Significance level			Panel Fisher test	
		1%	5%	10%	$\chi^2(2N)$	p -value
Engle–Granger	61	46	50	50	3223.4	< 0.001
Deviation unit root	61	47	48	49	3442.0	< 0.001

Both pair-by-pair tests examine each security in isolation. To evaluate the panel jointly, we combine the pair-level p -values using the Fisher test of Maddala and Wu (1999). Under the joint null hypothesis that every pair contains a unit root, the individual p -values are

⁷The Engle and Granger (1987) procedure has two weaknesses. Its power is low in samples of our length, because the ADF step cannot easily separate a persistent but stationary spread from a true unit root. In addition, its test result is sensitive to normalization. Regressing $\ln P$ on $\ln P^x$ yields slightly different, yet qualitatively similar results. At the 5% significance level, the residual-based test rejects the null of no cointegration for 47 of the 61 pairs.

independent and uniform, so that

$$-2 \sum_{i=1}^N \ln(p_i) \sim \chi^2(2N), \quad N = 61 \quad (2)$$

where p_i is the p -value of the corresponding test statistics. Table 2 reports the panel Fisher test results in the last two columns. The results reject the null hypothesis of no cointegration for the panel with p -values < 0.001 , suggesting that the token and underlying markets are cointegrated. Consistent with the no-arbitrage condition, the pair shares a common long-run trend, and any price deviation between them is temporary.

Furthermore, we test whether the cointegration status is systematically related to observable pair characteristics. The concern is that if the pairs that fail to reject were concentrated among the thinnest and most illiquid names, the non-rejections might indicate that the no-arbitrage condition genuinely breaks down where arbitrage is most impeded, rather than reflecting the limited power of unit-root tests in an eleven-month sample. To examine this, we regress an indicator for rejecting the no-cointegration null at the 5% level—under the Engle and Granger (1987) test (1_{EG}) and under the spread-based test (1_{dev})—on standardized pair-level liquidity and volatility characteristics, using both a linear probability model and a logit model:

$$1_i^{coint} = c + Amihud_i + Amihud_i^x + \log Vol D_i + \log Vol D_i^x + \sigma_i + \sigma_i^x + e_i \quad (3)$$

where the characteristic measures are described in Section 4 and superscript x denotes the tokens. We standardize the independent variables to zero mean and unit variance, so coefficients are per one-standard-deviation change. We use MacKinnon and White (1985) heteroskedasticity-robust standard errors for the OLS specifications and maximum-likelihood standard errors for the logit specifications.

Table 3 reports the estimation results. Pair characteristics have limited ability to explain which pairs are cointegrated, except for trading volume. The estimated coefficients

on log dollar volume of the underlying are positive and is the only regressor significant across all four specifications. The result suggests that a deeper, more actively traded equity market raises the probability that the token-underlying pair is cointegrated. The remaining characteristics are largely statistically insignificant. The coefficient on return volatility of either leg is statistically insignificant from zero throughout, and although the logit coefficients on the underlying's Amihud ratio are large in magnitude, their standard errors are larger still—an instability that traces to the extreme right-skew of the illiquidity measure documented in Table 1 and that renders these particular estimates uninformative.

Moreover, the token's own illiquidity does not predict cointegration status. The token's Amihud ratio, $Amihud^x$ is statistically insignificant in every specification. Pairs with less liquid tokens are no less likely to be cointegrated than pairs with liquid ones. The coefficients on token dollar volume, $\log VolD^x$, are negative and significant in the two linear probability models. Less traded tokens are more likely to be cointegrated. This result goes in line with the positive relationship of the trading volume in the underlying market. A more actively traded underlying market, as opposed to the token market, contributes more to the cointegration of the pair.

Taken together, the cointegration result places tokenized equities much closer to the well-arbitraged backed traditional securities markets than to the fragmentarily priced crypto markets. Cointegration for roughly three-quarters of pairs—and for the panel decisively—echoes the evidence on cross-listed securities and exchange-traded funds, where prices of the same claim in two markets share a common trend (Eun and Sabherwal, 2003; Engle and Sarkar, 2006), and contrasts with the large, persistent cross-venue price gaps documented for cryptocurrencies (Makarov and Schoar, 2020). Our results suggest this one-to-one backing is sufficient to enforce the no-arbitrage condition even in a market an order of magnitude thinner than the one it mirrors.

Table 3. Determinants of cointegration status

This table reports cross-sectional regressions relating each pair's cointegration status to its liquidity and volatility characteristics. The dependent variable in columns (1) and (2) is an indicator equal to one if the pair rejects the null of no cointegration in the Engle and Granger (1987) residual-based test at the 5% level (1^{EG}); in columns (3) and (4) it is an indicator equal to one if the pair rejects the unit-root null for the log price deviation $dev_t = \ln P_t - \ln P_t^x$ at the 5% level (1^{dev}), following Taylor and Taylor (2004). Columns (1) and (3) report linear probability (OLS) estimates; columns (2) and (4) report logit estimates. The independent variables are pair-level averages over the sample period: the Amihud (2002) illiquidity ratio of the underlying (Amihud) and of the token (Amihud^x), log dollar volume of each leg (logVolD, logVolD^x), and the return volatility of each leg (σ , σ^x). All independent variables are standardized to zero mean and unit variance. MacKinnon and White (1985) heteroskedasticity-robust standard errors are in parentheses for the OLS specifications; maximum-likelihood standard errors for the logit specifications. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

	1^{EG}		1^{dev}	
	(1)	(2)	(3)	(4)
Amihud	0.029 (0.847)	29.459 (92.755)	0.036 (0.906)	30.352 (106.852)
Amihud _x	0.051 (0.201)	23.352 (28.958)	0.059 (0.208)	29.246 (33.510)
logVolD	0.220** (0.109)	2.219** (0.986)	0.253** (0.113)	2.320** (0.939)
logVolD _x	-0.118** (0.058)	-0.431 (0.845)	-0.142** (0.067)	-0.697 (0.795)
σ	0.085 (0.089)	-1.419 (2.693)	0.126 (0.162)	0.065 (2.255)
σ_x	0.078 (0.070)	2.946 (2.782)	0.055 (0.160)	1.308 (2.277)
Constant	0.820*** (0.111)	10.953 (15.530)	0.787*** (0.119)	11.839 (17.899)
Estimation	OLS	Logit	OLS	Logit
Observations	61	61	61	61
(Pseudo) R^2	0.192	0.305	0.215	0.311

6 Price Discovery

The above cointegration result establishes that a token and its underlying equity are tethered to a common long-run value. Yet, it is silent on which market does the tethering. When new information arrives, one price could move first and the other follows, and when a gap opens between them, the markets would would the prices to closes it. Which market predominantly sets the market price and which one does most of the price adjustment depend on the price discovery process. The market where price discovery takes place would set the long term common price, and the other market would adjust to such price. In this section, we first estimate a vector error correction model for each cointegrated pair, whose adjustment coefficients reveal which price corrects disequilibrium. We then summarize the estimates with two standard decompositions: the permanent–transitory shares of Gonzalo and Granger (1995) and the information shares of Hasbrouck (1995).

Throughout, the timing of the data works against finding that the equity leads, for two reasons. First, the token close is stamped several hours after the equity close, so it mechanically embeds more same-day information. Second, the token trades continuously—nights, weekends, and holidays included—so between any two daily observations the token has a strictly larger window to incorporate news, much of which arrives while the equity market is closed. Both features tilt the daily data toward attributing price discovery to the token, making any equity-led result a conservative one.

6.1 Vector error correction model

First, we estimate the vector error correction models to capture the price adjustment to the long-run equilibrium price for each of the token and underlying markets. More specifically,

for each cointegrated pair we estimate the bivariate VECM:

$$\Delta \ln P_t^x = \alpha_x z_{t-1} + \sum_{i=1}^q \left(\Gamma_i^{(11)} \Delta \ln P_{t-i}^x + \Gamma_i^{(12)} \Delta \ln P_{t-i} \right) + \varepsilon_{x,t}, \quad (4)$$

$$\Delta \ln P_t = \alpha_u z_{t-1} + \sum_{i=1}^q \left(\Gamma_i^{(21)} \Delta \ln P_{t-i}^x + \Gamma_i^{(22)} \Delta \ln P_{t-i} \right) + \varepsilon_{u,t}, \quad (5)$$

$$z_{t-1} = \ln P_{t-1}^x - \ln P_{t-1} + c \quad (6)$$

with the cointegrating rank fixed at one and the lag length q chosen by the Akaike Information Criterion (AIC). The error-correction term z_{t-1} measures the disequilibrium between the two prices at the end of the previous trading day. The adjustment coefficients α_x and α_u measure how the token price and the underlying price respond to the disequilibrium, respectively. Their signs and magnitudes are of economic interests. A negative α_x means that when the token is overpriced relative to the underlying ($z_{t-1} > 0$), the token price falls back toward the underlying—the token does the price correction. A near zero α_u signals that the underlying barely responds to disequilibrium, so it behaves as the anchor around which the token adjusts. The long-run coefficient β_u describes the equilibrium relationship between the two prices. A value close to -1 indicates one-for-one comovement that would be consistent with a $(1, -1)$ spread.

Table 4 summarizes the coefficient estimates across the 51 cointegrated pairs.⁸ The individual pair estimation results are reported in Table A2 in the Appendix. The token's adjustment coefficient α_x has a mean of -0.602, and it is statistically significant at the 5% level in 72.6% of the pairs. The median is -0.575. The magnitude is also economically significant. An α_x coefficient near -0.6 implies that more than half of a given day's price

⁸The VECM estimation sample comprises 51 pairs, whereas the Engle and Granger (1987) test of Section 5 rejects the null of no cointegration for 50. The small discrepancy arises because the estimation sample is selected by the Johansen trace test, under which the VECM is estimated: the rank test and the error correction estimation derive from the same maximum-likelihood framework, making it the natural selection criterion at this stage. The Johansen procedure treats the two prices symmetrically and differs from the residual-based test in specification and power, so pairs near the 5% threshold can be classified differently under the two tests.

deviation between the token and its underlying prices is eliminated by the token’s price movement within the next trading day. The underlying’s adjustment coefficient α_u also supports this finding. Its mean of 0.228 carries a positive sign, and it is significant in only 27.5% of the pairs. In the typical pair, thus, the token does the majority of the price correction and the underlying responds less. When the two prices diverge, it is the token that moves to resolve the discrepancy. By the logic developed above, the underlying equity price serves as the anchor around which the common price forms.

Table 4. VECM adjustment and long-run coefficients

This table summarizes the coefficient estimates from the bivariate vector error correction model (VECM) of equations (4)–(6), estimated separately for each of the 51 cointegrated token–underlying pairs, with the cointegrating rank fixed at one and the lag length selected by the Akaike information criterion. α_x is the adjustment coefficient of the tokenized price and α_u that of the underlying price. The table reports the number of pairs, the cross-pair mean, the share of pairs in which the coefficient is statistically significant at the 5% level, and the quartiles of the cross-pair distribution. Significance is not reported for β_u because the cointegrating vector is not tested against zero.

	Count	Mean	% sig. at 5%	25th pct.	Median	75th pct.
α_x (token)	51	-0.602	72.55%	-0.865	-0.575	-0.398
α_u (underlying)	51	0.228	27.45%	-0.008	0.130	0.431
β_u (long run)	51	-0.998	-	-1.007	-0.999	-0.987

The long-run coefficient completes the picture. Across pairs, β_u is tightly concentrated around -1, with the interquartile range runs from -1.007 to -0.987 with a median of -0.999. So, the equilibrium relationship is one-for-one in log prices. In the long run, a one percent move in the underlying price is matched by a one percent move in the token price. This is the unit elasticity that the token’s one-for-one backing predicts.

6.2 Permanent–transitory decomposition

Building upon the VECM results, we investigate how the adjustment coefficients jointly determine the common price of each token-underlying pair—how much of it is set in the equity market and how much in the token market—by carrying out the permanent–transitory decomposition of Gonzalo and Granger (1995). This approach converts each pair’s estimated adjustment speeds into a single share of price discovery, with a sampling

distribution that allows the share to be tested pair by pair.

Specifically, Gonzalo and Granger (1995) identify the permanent component as the combination of the two prices that does not respond to the transitory part. The common trend is built from the markets in inverse proportion to how much they adjust. A market that does not move when a gap opens contributes everything to the common price, whereas a market that does all of the correcting contributes nothing, because its movements are the transitory noise being eliminated. Based on the VECM results above, the underlying's share of the common factor is

$$GG = \frac{-\alpha_x}{\alpha_u - \alpha_x}, \quad \alpha_u \neq \alpha_x \quad (7)$$

This measure maps the two adjustment speeds into a single share with natural interpretation: $GG = 1$ when the underlying equity price does not adjust at all ($\alpha_u = 0$), so that price discovery resides entirely in the equity; and $GG = 0$ when the token does not adjust ($\alpha_x = 0$), so that it resides entirely in the token; and $GG > 0.5$ when the equity contributes the majority of the price discovery, and the token does most of the correction ($-\alpha_x > \alpha_u$).

Table 5 reports the Gonzalo-Granger estimation results, and Panel A reports the cross-pair distribution of the estimates. The individual pair estimation results are reported in Table A3 in the Appendix. The result shows that, for the typical pair, the median is 0.825, suggesting that for the median pair more than 80% of the permanent price movement originates in the underlying and only 20% in the token. The cross-pair mean, 0.766, sits below the median because a handful of pairs carry an estimate above one—the 75th percentile is 1.009—which arise mechanically when an adjustment coefficient takes an unexpected sign rather than signaling more-than-total equity discovery. Winsorizing the results to the unit interval pulls the mean to 0.707 but leaves the median unchanged at 0.825. In other words, the underlying equity contributes the majority of price discovery.

Because GG is a smooth function of the estimated VECM adjustment coefficients, its sampling variance follows the Delta method. Writing $GG = g(\alpha_x, \alpha_u) = -\alpha_x/(\alpha_u - \alpha_x)$

Table 5. Gonzalo–Granger price-discovery shares

This table summarizes the Gonzalo and Granger (1995) permanent–transitory decomposition applied to each of the 51 cointegrated token–underlying pairs. For each pair, $GG = -\alpha_x/(\alpha_u - \alpha_x)$ is the share of the common permanent component attributable to the underlying equity, computed from the adjustment coefficients of the vector error correction model in Section 6.1. Panel A reports the cross-pair distribution of the raw share and of a version winsorized to $[0, 1]$. Panel B classifies pairs by whether GG differs from the even-split benchmark of 0.5 at the 5% level, using delta-method standard errors: the equity leads when GG is significantly above 0.5, the token leads when it is significantly below, and the split is indeterminate when the difference is insignificant.

Panel A: Distribution of price-discovery shares

	Count	Mean	Std.	25th pct.	Median	75th pct.
Gonzalo–Granger	51	0.766	0.427	0.520	0.825	1.009
Winsorized	51	0.707	0.336	0.520	0.825	1.000

Panel B: Leadership classification

Sig. at 5%	Count	Percentage
Equity leads ($GG > 0.5$)	21	41.18%
Indeterminate ($GG = 0.5$)	26	50.98%
Token leads ($GG < 0.5$)	4	7.84%

and letting $\hat{\Sigma}_\alpha$ denote the estimated covariance matrix of $(\hat{\alpha}_x, \hat{\alpha}_u)$, a first-order expansion gives

$$\widehat{\text{Var}}(\widehat{GG}) = \nabla g' \hat{\Sigma}_\alpha \nabla g, \quad \nabla g = \frac{1}{(\hat{\alpha}_u - \hat{\alpha}_x)^2} \begin{pmatrix} -\hat{\alpha}_u \\ \hat{\alpha}_x \end{pmatrix}, \quad (8)$$

where $\hat{\Sigma}_\alpha$ includes the cross-equation covariance of the two adjustment coefficients, which is nonzero because the two VECM equations share regressors and correlated innovations.

The classification in Panel B follows from a Wald test of the even-split benchmark: $t_{GG} = (\widehat{GG} - 0.5)/\sqrt{\widehat{\text{Var}}(\widehat{GG})}$ is asymptotically standard normal under the null $GG = 0.5$, and a pair is equity-led when $t_{GG} > 1.96$, token-led when $t_{GG} < -1.96$, and indeterminate otherwise. The result shows that, in 21 pairs (41.2%) the underlying equity leads in price discovery, and in only 4 pairs (7.8%) the token leads. For the remaining 26 pairs (51.0%) the split is statistically indeterminate. Worth noting that the indeterminate category captures the sample pairs that are imprecisely estimated to distinguish from an even split, and is accordingly concentrated among pairs with small adjustment gaps, where the Delta-method approximation is also weakest.⁹ Both features bias the procedure to-

⁹The squared adjustment gap in the denominator of the gradient (Eq. (8)) implies that shares near the

ward indeterminacy rather than toward either leadership claim, so the 21 equity-led and 4 token-led assignments survive a test that is conservative.

6.3 Information shares

While the Gonzalo–Granger decomposition shows that the tokenized market moves to restore equilibrium, we further address which market’s own shocks permanently move the common price. A market could adjust sluggishly to disequilibrium—behaving like a price leader under Gonzalo–Granger—while contributing little genuinely new information, or a market could impound the bulk of fundamental innovations even while also doing some of the correcting. To measure how the variance of new information is apportioned between the two markets, we estimate the information shares of Hasbrouck (1995), which decompose the innovation variance of the common efficient price into the contribution of each market.¹⁰

Hasbrouck’s measure attributes to each market the fraction of the efficient price’s innovation variance that originates in that market’s own price shocks. A market with a large information share is one whose shocks tend to become permanent features of the common price rather than transitory oscillation. The measure carries one complication. When the two markets’ price innovations are contemporaneously correlated (e.g., prices sampled at

even split are estimated least precisely.

¹⁰The information share is computed from the estimated VECM as follows. Inverting the vector error correction representation of equations (4)–(6) yields the vector moving-average form of the price changes, $\Delta p_t = \Psi(L) \varepsilon_t$, where $p_t = (\ln P_t^x, \ln P_t)^t$ and ε_t is the vector of reduced-form innovations with covariance matrix Ω . Let $\psi = (\psi_x, \psi_u)$ denote the row vector $\Psi(1)$, common to both equations because the two prices share a single stochastic trend; its entries measure the long-run impact of each market’s innovation on the common efficient price. The variance of the increment to the efficient price is then $\psi \Omega \psi'$. When Ω is diagonal—the innovations are contemporaneously uncorrelated—market j ’s information share is simply its contribution to this variance, $IS_j = \psi_j^2 \Omega_{jj} / (\psi \Omega \psi')$. When Ω is not diagonal, the correlated component of the innovations must be attributed to one market or the other. Following Hasbrouck (1995), we eliminate the contemporaneous correlation with the Cholesky factorization $\Omega = FF'$, where F is lower triangular, and compute $IS_j = ([\psi F]_j)^2 / (\psi \Omega \psi')$. Because F depends on the order in which the two prices are placed, the market ordered first is credited with the entire shared innovation, and the share is therefore not unique. We compute the equity’s share under both orderings: placing the equity first yields its upper bound and placing it second its lower bound. All shares are estimated pair by pair, using the AIC-selected lag length from the corresponding VECM.

the same daily frequency) the decomposition is not unique. It depends on the order in which the markets are placed in the Cholesky factorization of the innovation covariance matrix. Rather than impose an arbitrary ordering, we follow standard practice and report the full range, computing the equity's share under both orderings to obtain a lower and an upper bound, along with their midpoint. The tighter these bounds, the less the ordering matters and the more informative the measure.

Table 6 reports the results, which corroborate the Gonzalo–Granger evidence. The equity's information share is centered above 0.5. The midpoint has a mean of 0.618 and a median of 0.611. So the two decompositions built on different criteria—adjustment speeds and innovation variances—place the equity's contribution in the same place. The bounds themselves are also informative about where each market stands. The equity's upper bound is close to one throughout the distribution (mean 0.947, median 0.980), while its lower bound has a mean of 0.288 and a median of 0.240.

Table 6. Hasbrouck information shares

This table summarizes the Hasbrouck (1995) information shares of the underlying equity for each of the 51 cointegrated token–underlying pairs. For each pair we report a lower bound, an upper bound, and their midpoint, depending on the ordering of the two markets in the Cholesky factorization of the innovation covariance matrix. Values are expressed as the equity's share. Panel A reports the cross-pair distribution of the lower bound, upper bound, and midpoint. Panel B classifies pairs by leadership at the 0.5 benchmark: the equity leads when its lower bound exceeds 0.5, the token leads when the equity's upper bound falls below 0.5, and the assignment is inconclusive when the bounds straddle 0.5.

Panel A: Distribution of information shares

	Count	Mean	Std.	25th pct.	Median	75th pct.
Lower bound	51	0.288	0.277	0.045	0.240	0.454
Upper bound	51	0.947	0.080	0.938	0.980	0.996
Midpoint	51	0.618	0.158	0.499	0.611	0.721

Panel B: Leadership classification at the 0.5 benchmark

Test	Count	Percentage
Equity leads (lower bound > 0.5)	11	21.57%
Bounds straddle 0.5	40	78.43%
Token leads (upper bound < 0.5)	0	0.00%

The leadership classification in Panel B shows the asymmetry of information share be-

tween the tokenized equity and the underlying equity. The result shows that the equity market contributes more to the price discovery, as the information shares' lower bound alone exceeds 0.5 in 11 of 51 pairs. Whereas the token leads in none, as in not a single pair does the equity's upper bound fall below 0.5. The remaining 40 pairs have bounds that straddle 0.5, which is the expected consequence of wide daily bounds.

Overall, our findings extend to tokenized equities a regularity long established for other multi-market claims—that price discovery concentrates in the more liquid, established venue—and place them at the strong end of that range (Hasbrouck, 1995; Eun and Sabherwal, 2003; Hasbrouck, 2003). The leading venue is a regulated equity market and the following venue is a around-the-clock trading, on-chain market connected to the underlying only by a redeemable custodial claim. That this claim suffices to keep essentially all price discovery in the underlying, in contrast to the fragmented, leaderless pricing documented across cryptocurrency venues (Makarov and Schoar, 2020).

7 Determinants of Price Discovery

The evidence above shows that the underlying equity market tends to lead and contribute more to the price discovery process of the token-underlying pair. In this section, we investigate what accounts for the variation in the price discovery across different pairs.

Specifically, we regress each pair's price-discovery measures on its liquidity and volatility characteristics. The dependent variables include (1) the Gonzalo–Granger share and (2) its winsorized version, (3) the three-way leadership classification from Panel B of Table 5 (estimated by ordered logit), (4) the Hasbrouck (1995) information share midpoint, and an indicator for equity leadership by the information share criterion from Panel B of Table 6 (estimated by logit). The regressors are pair-level averages of the Amihud (2002) illiquidity ratio for each leg, the token's share of the pair's combined dollar volume, and each leg's return volatility. The OLS models (Columns (1), (2), and (4)) are estimated with

MacKinnon and White (1985) heteroskedasticity-robust standard errors. Table 7 reports the results.

Table 7. Determinants of price discovery

This table reports cross-sectional regressions relating each pair's price-discovery measures to its liquidity and volatility characteristics across the 51 cointegrated token-underlying pairs. The dependent variable in column (1) is the Gonzalo and Granger (1995) share GG of the underlying equity; in column (2) its version winsorized to the unit interval; in column (3) a categorical leadership indicator taking the value 1 when the equity leads (GG significantly above 0.5 at the 5% level), -1 when the token leads (GG significantly below 0.5), and 0 when leadership is indeterminate, estimated by ordered logit; in column (4) the midpoint of the Hasbrouck (1995) information-share bounds; and in column (5) an indicator equal to one when the equity leads by the information-share criterion—its lower bound exceeds 0.5—estimated by logit. Columns (1), (2), and (4) report OLS estimates with MacKinnon and White (1985) heteroskedasticity-robust standard errors. Regressors are pair-level averages over the sample period: the Amihud (2002) illiquidity ratio of the underlying ($Amihud$) and of the token ($Amihud_x$), the token's share of the pair's combined dollar volume ($VolD_{pct}$), and the return volatility of each leg (σ , σ_x). The rows $\mu_{-1|0}$ and $\mu_{0|1}$ report the estimated cutpoints of the ordered logit. Standard errors are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

	GG (1)	GG (wins.) (2)	$Leads_{GG}$ (3)	IS midpoint (4)	$Leads_{IS}$ (5)
Amihud	0.150 (0.849)	0.049 (0.456)	5.491 (6.995)	0.036 (0.362)	2.056 (20.486)
Amihud _x	0.143 (0.273)	0.102 (0.153)	-0.870 (1.441)	0.062 (0.076)	7.051 (5.582)
VolD _{pct}	-0.111 (0.096)	-0.117* (0.067)	-0.295 (0.515)	-0.051** (0.021)	-0.050 (0.838)
σ	-0.316 (0.404)	-0.275 (0.256)	-5.755** (2.471)	-0.164** (0.070)	-12.672** (6.134)
σ_x	0.180 (0.420)	0.166 (0.261)	5.453** (2.516)	0.115* (0.064)	7.011 (4.509)
Constant	0.766*** (0.126)	0.707*** (0.074)		0.618*** (0.050)	-3.924 (3.815)
$\mu_{-1 0}$			-4.486*** (1.304)		
$\mu_{0 1}$			1.422*** (0.201)		
Estimation	OLS	OLS	Ordered logit	OLS	Logit
Observations	51	51	51	51	51
(Pseudo) R^2	0.430	0.414	0.281	0.547	0.579

The result shows that the division of price discovery across pairs is governed mainly by the relative volatility of the two legs rather than by the level of token illiquidity. A more volatile underlying lowers the equity's share of discovery, while a more volatile token raises it. The coefficients on σ and σ_x carry opposite signs in every column and attain significance

in the leadership classification of column (3) and the information-share specifications of columns (4) and (5). The raw Gonzalo–Granger level in columns (1) and (2) traces the same directional pattern more weakly.

Economically, the finding suggests that the decompositions attribute discovery away from the market whose movements are transitory. Return volatility beyond what fundamental news warrants is the disequilibrium that a market subsequently corrects. The leg that diverges is doing the price correction rather than contributing to the price discovery, and is assigned a smaller permanent share as a result. Crucially, own-leg volatility lowers a market’s own discovery share rather than raising it—the reverse of what one would see if volatility proxied for the pace of information arrival—which identifies these coefficients as capturing transitory, correctable movement rather than fundamental innovation. A volatile token therefore raises GG , while noise in the underlying works in the opposite direction.

Token illiquidity, by contrast, does not robustly explain the division of discovery. The coefficient on the token’s Amihud ratio, $Amihud_x$, keeps the positive sign and roughly the magnitude it takes in tighter specifications, but its standard error is too large for it to be distinguished from zero in any column on the 51 cointegrated pairs. The underlying’s own illiquidity is also insignificant throughout. The one liquidity characteristic that does register is the token’s share of the pair’s combined dollar volume, $VolD_{pct}$, which enters negatively in the winsorized- GG and information-share columns: where the token accounts for a larger fraction of trading, it captures a larger fraction of discovery, lowering the equity’s share. This is the relative-activity mechanism documented across venues by Frijns, Gilbert, and Tourani-Rad (2015), under which a market’s contribution to price discovery rises with its relative trading activity, though here the effect is only marginally significant.

Taken together, the determinants place the source of cross-pair variation in the relative amount of transitory, correctable movement on each leg rather than in the cost of arbitraging the token. Significance is concentrated in the leadership and information-share specifications; the pair-by-pair GG level, though measured with more precision, yields a

directionally identical but statistically weaker picture, consistent with the loss of power from regressing a noisy continuous share on only six standardized characteristics across 51 pairs.

8 Conclusion

Tokenized equities promise the exposure of a listed stock in the wrapper of a blockchain-based asset, that is self-custodial, transferable across platforms, and tradable around the clock. Whether the wrapper preserves the economics of the underlying claim is an empirical question. Using a daily panel of 61 matched token–underlying pairs over the market’s first eleven months, we investigate whether the token and underlying equity prices are tied together in the long run and, if so, which market sets the common price.

First, we find that the tether holds. Both legs’ prices are cointegrated of order one for roughly 80% of pairs under both the residual-based test of Engle and Granger (1987) and a deviation unit-root test imposing the one-for-one backing, and the long-run relationship is almost exactly one-for-one in log prices. The no-arbitrage condition thus survives transplantation into an on-chain venue an order of magnitude thinner than the market it mirrors. Secondly the formation of common price is anchored to the traditional securities market. The token does the bulk of the adjustment when price deviation takes place, as its adjustment coefficient averages 0.602 and is significant in nearly three-quarters of pairs, whereas the underlying barely responds. The Gonzalo and Granger (1995) decomposition attributes more than 80% of the permanent component to the equity for the median pair, and the Hasbrouck (1995) information shares corroborate the finding. The equity leads in about 20% of pairs by the most conservative criterion, and the token leads in none. Tokenized equities therefore follow the regularity long documented for cross-listed shares and index products (Eun and Sabherwal, 2003; Hasbrouck, 2003) that the deeper, established venue leads, and the satellite adjusts.

The result carries implications for exchanges and regulators. For exchanges, it indicates that main trading venue forms the price discovery, which does not migrate merely because an always-open alternative exists. Tokenization as currently constituted is not disintermediation but a distribution channel, extending the exchange's prices to new hours and investors while contributing little to their price discovery. For regulators, the finding admits a reassuring reading, in that the informational integrity of token prices rests on a market inside the existing perimeter. For regulators, the finding admits a reassuring reading, in that the informational integrity of token prices rests on a market inside the existing perimeter. But the token market inherits those prices, and the issuer is in effect distributing the regulated market's prices without the regulated market's obligations, so investors obtain the primary market's prices without its protections. The regulated equity market leads price discovery does not neutralize the need for regulatory oversight.

Lastly, our analysis is subject to limits that mark out the next steps. Prices are sampled once per day, so the paper is silent on intraday dynamics; the sample covers the market's first eleven months, when token volumes were a small fraction of underlying volumes; and the setting comprises a single issuer and token venue. As the market deepens, higher-frequency data spanning exchange and off-exchange hours would allow the intraday and overnight dimensions of discovery—the latter studied by Cong et al. (2025)—to be integrated with the cross-market dimension examined here. A further extension follows from a change in market structure now underway. The U.S. SEC has approved moves by U.S. exchanges to 23-hour, five-day trading. Because the extension alters the underlying's trading hours while leaving the tokens' 24/7 schedule untouched, comparing price discovery before and after would separate the contribution of trading hours from that of order flow.

References

- Ackert, L. F., and Y. S. Tian. 2000. Arbitrage and valuation in the market for standard and poor's depositary receipts. *Financial Management* 71–87.
- Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5:31–56.
- Ben-David, I., F. Franzoni, and R. Moussawi. 2018. Do ETFs increase volatility? *The Journal of Finance* 73:2471–535.
- Brandvold, M., P. Molnár, K. Vagstad, and O. C. A. Valstad. 2015. Price discovery on bitcoin exchanges. *Journal of International Financial Markets, Institutions and Money* 36:18–35.
- Chakravarty, S., H. Gulen, and S. Mayhew. 2004. Informed trading in stock and option markets. *The Journal of Finance* 59:1235–57.
- Cong, L. W., W. R. Landsman, D. Rabetti, C. Zhang, and W. Zhao. 2025. Tokenized stocks. Available at SSRN 5937314 .
- Engle, R., and D. Sarkar. 2006. Premiums-discounts and exchange traded funds. *Journal of Derivatives* 13:27–.
- Engle, R. F., and C. W. Granger. 1987. Co-integration and error correction: Representation, estimation, and testing. *Econometrica: Journal of the Econometric Society* 251–76.
- Eun, C. S., and S. Sabherwal. 2003. Cross-border listings and price discovery: Evidence from US-listed Canadian stocks. *The Journal of Finance* 58:549–75.
- Frijns, B., A. Gilbert, and A. Tourani-Rad. 2015. The determinants of price discovery: Evidence from US-Canadian cross-listed shares. *Journal of Banking & Finance* 59:457–68.

- Garbade, K. D., and W. L. Silber. 1983. Price movements and price discovery in futures and cash markets. *The Review of Economics and Statistics* 289–97.
- Gonzalo, J., and C. Granger. 1995. Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics* 13:27–35.
- Grammig, J., M. Melvin, and C. Schlag. 2005. Internationally cross-listed stock prices during overlapping trading hours: Price discovery and exchange rate effects. *Journal of Empirical Finance* 12:139–64.
- Griffin, J. M., and A. Shams. 2020. Is bitcoin really untethered? *The Journal of Finance* 75:1913–64.
- Hasbrouck, J. 1995. One security, many markets: Determining the contributions to price discovery. *The Journal of Finance* 50:1175–99.
- . 2003. Intraday price formation in US equity index markets. *The Journal of Finance* 58:2375–400.
- Hautsch, N., C. Scheuch, and S. Voigt. 2024. Building trust takes time: Limits to arbitrage for blockchain-based assets. *Review of Finance* .
- Lin, K. 2026a. Blockchain settlement latency, market liquidity, and arbitrage in cryptocurrency markets. *WFE Research Working Paper, Proceedings of the EUROFIDAI-ESSEC Paris December Finance Meeting 2024*. Available at SSRN: <https://ssrn.com/abstract=4807552> .
- . 2026b. Extending exchange trading hours: A coordination perspective. *WFE Research Working Paper*, Available at SSRN: <https://ssrn.com/abstract=6224760> or <http://dx.doi.org/10.2139/ssrn.6224760> .
- Lyons, R. K., and G. Viswanath-Natraj. 2023. What keeps stablecoins stable? *Journal of International Money and Finance* 131:102777–.

- MacKinnon, J. G., and H. White. 1985. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics* 29:305–25.
- Maddala, G. S., and S. Wu. 1999. A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics* 61:631–52.
- Makarov, I., and A. Schoar. 2020. Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics* 135:293–319.
- Menkveld, A. J. 2008. Splitting orders in overlapping markets: A study of cross-listed stocks. *Journal of Financial Intermediation* 17:145–74.
- Petajisto, A. 2017. Inefficiencies in the pricing of exchange-traded funds. *Financial Analysts Journal* 73:24–54.
- Taylor, A. M., and M. P. Taylor. 2004. The purchasing power parity debate. *Journal of Economic Perspectives* 18:135–58.
- Werner, I. M., and A. W. Kleidon. 1996. UK and US trading of British cross-listed stocks: An intraday analysis of market integration. *The Review of Financial Studies* 9:619–64.

Appendix A. Additional Results

Table A1. Augmented Dickey–Fuller unit-root tests

This table reports p -values from augmented Dickey–Fuller (ADF) tests of the null hypothesis of a unit root, applied to the daily log closing price of each security in levels, $\ln(\text{Price})$, and in first differences, $\Delta \ln(\text{Price})$. For each matched pair, the test is run separately on the underlying U.S.-listed security and on its tokenized counterpart. N is the number of daily observations available for the pair.

Ticker	N	$\ln(\text{Price})$		$\Delta \ln(\text{Price})$	
		Underlying	Token	Underlying	Token
AAPL	230	0.717	0.744	0.000	0.000
ABBV	175	0.241	0.024	0.000	0.000
ABT	175	0.969	0.888	0.000	0.000
ACN	175	0.855	0.883	0.000	0.000
AMBR	227	0.131	0.200	0.000	0.000
AMZN	227	0.465	0.482	0.000	0.000
APP	150	0.514	0.523	0.000	0.000
AVGO	175	0.685	0.560	0.000	0.000
AZN	93	0.350	0.344	0.000	0.000
BAC	150	0.272	0.269	0.000	0.001
BRK.B	180	0.108	0.242	0.000	0.000
CMCSA	175	0.477	0.249	0.000	0.000
COIN	230	0.705	0.714	0.000	0.000
CRCL	230	0.311	0.312	0.000	0.000
CRM	150	0.523	0.576	0.189	0.000
CRWD	175	0.982	0.943	0.000	0.000
CSCO	175	0.995	0.995	0.000	0.000
CVX	150	0.729	0.713	0.000	0.000
DFDV	226	0.673	0.682	0.000	0.000

Table A1 (continued)

Ticker	N	$\ln(\text{Price})$		$\Delta \ln(\text{Price})$	
		Underlying	Token	Underlying	Token
DHR	175	0.782	0.759	0.000	0.000
GLD	220	0.436	0.531	0.000	0.000
GME	155	0.173	0.027	0.000	0.000
GOOGL	230	0.481	0.519	0.000	0.000
HD	175	0.457	0.427	0.000	0.000
HON	175	0.745	0.725	0.000	0.000
HOOD	228	0.580	0.582	0.000	0.000
IBM	150	0.608	0.544	0.000	0.000
INTC	150	0.985	0.989	0.000	0.000
JNJ	150	0.202	0.418	0.000	0.000
JPM	150	0.081	0.099	0.000	0.000
KO	175	0.598	0.598	0.000	0.000
LIN	150	0.843	0.900	0.000	0.000
LLY	174	0.225	0.167	0.000	0.000
MA	174	0.383	0.403	0.000	0.000
MCD	224	0.649	0.616	0.000	0.000
MDT	150	0.990	0.945	0.000	0.000
META	230	0.297	0.282	0.000	0.000
MRK	150	0.287	0.297	0.000	0.000
MRVL	150	0.994	0.982	0.000	0.000
MSFT	168	0.510	0.542	0.000	0.000
MSTR	230	0.626	0.627	0.000	0.000
NFLX	160	0.465	0.196	0.000	0.000

Table A1 (continued)

Ticker	N	$\ln(\text{Price})$		$\Delta \ln(\text{Price})$	
		Underlying	Token	Underlying	Token
NVDA	230	0.072	0.057	0.000	0.000
NVO	150	0.436	0.497	0.000	0.000
OPEN	150	0.373	0.260	0.000	0.000
ORCL	150	0.160	0.197	0.000	0.000
PEP	174	0.073	0.244	0.004	0.000
PFE	150	0.343	0.207	0.000	0.000
PG	174	0.209	0.405	0.000	0.177
PLTR	150	0.371	0.392	0.000	0.000
PM	150	0.558	0.495	0.000	0.000
QQQ	230	0.964	0.949	0.000	0.000
SPY	230	0.843	0.865	0.000	0.000
TMO	150	0.735	0.666	0.000	0.000
TONX	149	0.078	0.244	0.000	0.000
TQQQ	154	0.963	0.962	0.000	0.000
TSLA	230	0.172	0.177	0.000	0.000
UNH	192	0.592	0.032	0.000	0.000
V	150	0.213	0.243	0.000	0.000
VTI	174	0.904	0.969	0.000	0.000
XOM	150	0.631	0.608	0.000	0.000

Table A2. Vector error correction model estimates by ticker

This table reports the vector error correction model estimates in Equations (4)–(6) for each of the 51 cointegrated token–underlying pairs. The cointegrating rank is fixed at one and the lag length (lag) is selected by the Akaike information criterion; N is the number of daily observations. α_x and α_u are the adjustment coefficients of the tokenized and underlying prices, respectively, measuring the speed at which each price responds to the previous day’s deviation z_{t-1} between the two markets; each is followed by the p -value for the test that the coefficient equals zero. β_u is the long-run cointegrating coefficient in Equation (6).

Ticker	N	lag	α_x	p -value	α_u	p -value	β_u
AAPL	230	1	-0.237	0.432	0.788	0.006	-1.000
ABT	175	1	-0.890	0.000	-0.206	0.063	-0.967
AMBR	227	3	-0.251	0.000	-0.109	0.026	-0.916
AMZN	227	1	-0.130	0.451	0.971	0.000	-1.007
APP	150	1	-0.822	0.007	0.174	0.530	-1.013
AVGO	175	1	-0.644	0.004	0.241	0.246	-0.985
AZN	93	1	-0.867	0.014	-0.092	0.787	-1.016
BAC	150	1	-0.937	0.000	0.062	0.596	-1.006
BRK.B	180	1	-0.844	0.000	-0.008	0.919	-1.043
COIN	230	1	-0.349	0.418	0.640	0.126	-1.000
CRCL	230	1	-0.436	0.448	0.308	0.593	-0.996
CRM	150	1	-0.575	0.000	0.134	0.371	-1.001
CRWD	175	1	-0.691	0.000	-0.043	0.768	-0.992
CSCO	175	1	-0.414	0.022	0.434	0.000	-1.016
CVX	150	1	-0.670	0.000	0.006	0.951	-1.022
DFDV	226	1	0.079	0.812	0.945	0.007	-0.992
GLD	220	1	-0.469	0.002	-0.124	0.408	-0.993
GME	155	1	-0.324	0.000	0.047	0.478	-0.959
GOOGL	230	1	-0.132	0.537	0.854	0.000	-1.003
HOOD	228	1	0.113	0.753	1.122	0.001	-1.004
IBM	150	1	-0.381	0.044	0.298	0.083	-0.988

Table A2 (continued)

Ticker	N	lag	α_x	p-value	α_u	p-value	β_u
INTC	150	1	-0.011	0.963	0.982	0.000	-1.001
JNJ	150	1	-1.036	0.000	0.025	0.842	-0.986
JPM	150	1	-0.917	0.000	-0.023	0.850	-0.961
LIN	150	1	-0.913	0.000	-0.123	0.124	-0.995
LLY	174	1	-1.458	0.000	-0.480	0.036	-1.000
MA	174	1	-0.489	0.009	0.173	0.402	-0.976
MDT	150	1	-0.852	0.000	0.069	0.470	-0.985
META	230	1	-0.009	0.964	1.047	0.000	-1.001
MRK	150	1	-0.977	0.000	-0.007	0.950	-0.969
MRVL	150	1	-0.728	0.002	0.262	0.180	-0.999
MSFT	168	2	-0.558	0.000	0.120	0.341	-1.009
MSTR	230	1	-0.418	0.313	0.522	0.219	-0.999
NFLX	160	1	-0.747	0.000	0.074	0.138	-0.992
NVDA	230	1	-1.064	0.002	0.005	0.987	-1.004
NVO	150	1	-0.504	0.094	0.437	0.122	-1.007
OPEN	150	1	-1.359	0.000	-0.185	0.307	-0.998
ORCL	150	1	-0.569	0.019	0.495	0.014	-0.996
PEP	174	3	-0.153	0.301	0.264	0.033	-1.145
PFE	150	3	-0.978	0.000	-0.190	0.162	-0.924
PLTR	150	1	-1.313	0.000	-0.230	0.314	-0.984
PM	150	1	-0.616	0.000	0.100	0.426	-1.032
QQQ	230	1	-0.740	0.023	0.203	0.506	-1.003
TMO	150	1	-0.526	0.000	0.098	0.238	-0.996
TONX	149	2	-0.425	0.000	-0.116	0.338	-0.964

Table A2 (continued)

Ticker	N	lag	α_x	p -value	α_u	p -value	β_u
TQQQ	154	1	-0.721	0.039	0.133	0.690	-0.994
TSLA	230	1	-0.435	0.273	0.428	0.278	-1.000
UNH	192	1	0.024	0.883	0.713	0.000	-0.946
V	150	1	-0.863	0.000	0.206	0.023	-1.069
VTI	174	1	-0.494	0.000	0.130	0.243	-1.047
XOM	150	1	-0.967	0.000	0.045	0.753	-1.007

Table A3. Gonzalo–Granger price-discovery shares by ticker

This table reports the Gonzalo and Granger (1995) permanent–transitory decomposition for each of the 51 cointegrated token–underlying pairs. For each pair, the Gonzalo–Granger statistic $GG = -\alpha_x/(\alpha_u - \alpha_x)$ is the share of the common permanent component attributable to the underlying equity, computed from the adjustment coefficients of the vector error correction model in Section 6.1: $GG = 1$ indicates that price discovery resides entirely in the equity and $GG = 0$ entirely in the token. Standard errors are obtained by the delta method. N is the number of daily observations. The final column reports the share winsorized to the unit interval; values outside $[0,1]$, which arise when an adjustment coefficient carries an unexpected sign, are set to the nearer bound of 0 or 1.

Ticker	N	Gonzalo–Granger	Standard Error	Winsorized
AAPL	230	0.231	0.289	0.231
ABT	175	1.301	0.170	1.000
AMBR	227	1.763	0.493	1.000
AMZN	227	0.118	0.150	0.118
APP	150	0.825	0.281	0.825
AVGO	175	0.727	0.234	0.727
AZN	93	1.119	0.445	1.000
BAC	150	0.938	0.116	0.938
BRK.B	180	1.010	0.098	1.000
COIN	230	0.353	0.429	0.353
CRCL	230	0.586	0.771	0.586
CRM	150	0.811	0.205	0.811
CRWD	175	1.066	0.225	1.000
CSCO	175	0.488	0.167	0.488
CVX	150	0.991	0.150	0.991
DFDV	226	-0.091	0.383	0.000
GLD	220	1.360	0.445	1.000
GME	155	0.872	0.179	0.872
GOOGL	230	0.134	0.210	0.134
HOOD	228	-0.112	0.359	0.000
IBM	150	0.561	0.253	0.561

Table A3 (continued)

Ticker	N	Gonzalo–Granger	Standard Error	Winsorized
INTC	150	0.011	0.245	0.011
JNJ	150	0.976	0.120	0.976
JPM	150	1.026	0.139	1.000
LIN	150	1.156	0.111	1.000
LLY	174	1.490	0.246	1.000
MA	174	0.739	0.298	0.739
MDT	150	0.925	0.099	0.925
META	230	0.008	0.178	0.008
MRK	150	1.007	0.116	1.000
MRVL	150	0.736	0.202	0.736
MSFT	168	0.823	0.181	0.823
MSTR	230	0.445	0.443	0.445
NFLX	160	0.910	0.061	0.910
NVDA	230	0.995	0.308	0.995
NVO	150	0.536	0.304	0.536
OPEN	150	1.158	0.156	1.000
ORCL	150	0.535	0.200	0.535
PEP	174	0.367	0.315	0.367
PFE	150	1.241	0.180	1.000
PLTR	150	1.212	0.213	1.000
PM	150	0.860	0.173	0.860
QQQ	230	0.784	0.326	0.784
TMO	150	0.844	0.127	0.844
TONX	149	1.373	0.464	1.000

Table A3 (continued)

Ticker	N	Gonzalo–Granger	Standard Error	Winsorized
TQQQ	154	0.845	0.390	0.845
TSLA	230	0.504	0.456	0.504
UNH	192	-0.035	0.240	0.000
V	150	0.807	0.084	0.807
VTI	174	0.791	0.177	0.791
XOM	150	0.955	0.141	0.955