Non-Commuting Price Discovery in Fragmented Equity Markets:

Evidence from Multi-Venue IBM Prices

Johannes Bleher

June 17, 2026

Abstract

Price discovery in fragmented equity markets requires separating common efficient-price information from venue-specific adjustment. This paper studies that problem in a synchronized March 2021 panel of IBM prices from the five U.S. trading venues analyzed by Dias and Schweikert (2026): NYSE (N), Nasdaq (T), Arca (P), Cboe EDGX (K), and Cboe BZX (Z). We estimate cointegrated vector error-correction models (VECMs) for log prices sampled at 1, 5, 10, 30, and 60 seconds. The VECM benchmark has one permanent component at every frequency, but its transition matrices are not frequency consistent: Frobenius norms of the defects $F_{k\delta} - F_{\delta}^{k}$ range from 4.67 to 5.87 and are bounded away from zero by block-bootstrap intervals. Hasbrouck information-share bounds place N first at the one-second frequency, with midpoint share 0.474, followed by T and Z. Short-lag return correlations instead reveal directional timing asymmetries, especially from P, T, K, and Z toward N. We propose a commutator diagnostic for market-specific VECM update operators and use it to measure order-dependence in the fitted adjustment system. The largest beta-projected commutator variance is 1.215×10^{-11} for N/Z, equal to a combined spread-projection standard deviation of about 0.035 basis points, and projected effects exhibit excess kurtosis as high as 15.184. In this setting, fragmented price discovery is therefore not exhausted by a single permanent-price contribution ranking for IBM; it also contains frequency-dependent and non-commuting adjustment components.

Keywords: fragmented markets; price discovery; vector error-correction models; Hasbrouck information shares; commutator diagnostic; realized volatility.

JEL classification: C32; C58; G14.

1 Introduction

Price discovery in modern U.S. equities takes place across fragmented venues. Prices for the same security should be tied to a common efficient price, but at high frequency each venue also reflects its own trading frictions, clientele, liquidity, and market design. This paper studies which parts of this multi-venue system carry permanent-price information and which parts reflect venue-specific adjustment.

The statistical difficulty is that standard price-discovery tools compress a high-dimensional adjustment system into a small number of summaries. Cointegrated VECMs and Hasbrouck information shares are natural benchmarks because they impose one permanent price and stationary transitory spreads (Hasbrouck, 1995; Johansen, 1991). This structure is useful, but it does not directly test whether estimated adjustment dynamics are stable across sampling frequencies, nor whether market-specific adjustment channels commute with one another. A further practical difficulty is that calendar-time plots and returns can be misleading when overnight and weekend closures are treated as long real-time intervals rather than as breaks in trading activity.

This paper asks whether price discovery in fragmented markets is adequately described by a single permanent-price contribution ranking. To answer this question, we estimate a cointegrated benchmark for synchronized multi-venue prices, test whether the transition matrices are consistent across sampling frequencies, distinguish short-run lead–lag timing from Hasbrouck information shares, and introduce a commutator diagnostic for non-commuting market-update operators. The empirical object is the synchronized log-price vector

$$X_t = \begin{pmatrix} \log N_t & \log T_t & \log K_t & \log P_t & \log Z_t \end{pmatrix}', \qquad r_t = X_t - X_{t-1}.$$

The first contribution is empirical. In the March 2021 IBM sample from the five venues studied by Dias and Schweikert (2026), Johansen tests select cointegration rank four at every sampling frequency, leaving a single common permanent component. Yet the companion transitions are strongly frequency dependent: Frobenius norms of $F_{k\delta} - F_{\delta}^{k}$ range from 4.67 to 5.87, and all reported block-bootstrap intervals are bounded away from zero. Hasbrouck midpoint shares at the one-second frequency rank N, T, Z, K, and P, whereas one-second lead–lag correlations show that P, T, K, and Z all lead N in pairwise short-horizon return timing. The evidence therefore

separates permanent-price contribution from short-run timing.

The second contribution is methodological. From the fitted VECM transition matrix, we construct market-specific row-update operators and compare the two possible orderings of each market pair. If the two operators commute, the implied local transition is invariant to update order; if they do not, the fitted fragmented price system contains order-dependent adjustment. The diagnostic detects statistically visible but price-level-small non-commutation: the largest beta-projected commutator variance is 1.215×10^{-11} for N/Z, corresponding to about 0.035 basis points on a combined spread-projection standard-deviation scale, and projected effects have excess kurtosis as high as 15.184. Thus fragmented price discovery contains a component that is invisible to scalar permanent-price contribution measures.

The remainder of the paper is structured as follows. Section 2 positions the paper in the related literature. Section 3 describes the synchronized IBM data and the treatment of market-closure intervals. Section 4 introduces the cointegrated VECM benchmark. Section 5 tests frequency consistency. Section 6 compares lead–lag correlations with Hasbrouck information shares. Section 7 develops the commutator diagnostic, and Section 8 reports its empirical estimates. Section 9 studies volatility and higher moments. Section 10 concludes with practical implications. The Appendix reports bounded lag- and month-robustness screens and an illustrative monetizability screen.

2 Related Literature

The paper contributes first to the literature on fragmented markets and high-frequency market microstructure. O'Hara and Ye (2011) show that fragmentation need not mechanically harm market quality, while O'Hara (2015) and Menkveld (2016) emphasize that modern high-frequency markets operate through heterogeneous venues, trading protocols, and information channels. In this environment, venue-specific prices can differ even when the law of one price ties them to the same latent asset value. The empirical analysis uses precisely this institutional feature: the object of interest is not a single consolidated price, but a synchronized system of venue-level prices whose adjustment dynamics may differ across markets.

The second related strand is the literature on realized volatility and market microstructure noise. Classical realized-variance estimators are sensitive to high-frequency noise, leading to noise-robust

approaches such as realized kernels and pre-averaging (Hansen and Lunde, 2006; Bandi and Russell, 2008; Barndorff-Nielsen et al., 2008; Christensen et al., 2010). Dias and Schweikert (2026) extend this discussion by identifying fragmentation noise as a multiplicative noise component that arises because information about the efficient price is dispersed across venues. Their two-step approach uses multi-venue error-correction dynamics to recover the efficient component before estimating integrated variance. The present paper uses the same multi-venue logic, but asks whether the estimated adjustment dynamics are frequency-stable and whether venue-specific update channels commute.

The third strand is cointegration-based price discovery. Hasbrouck (1995) interprets permanent innovations in a cointegrated price system as information shares, and Johansen (1991) provides the likelihood framework for estimating the cointegration space. Subsequent work has adapted this logic to richer high-frequency settings, including continuous-time price discovery and cross-market information processing (Dias et al., 2021; Scherrer, 2021). The analysis below retains this benchmark but separates three objects that are often conflated in applied work: permanent-price contribution, short-run lead–lag timing, and order-dependence of the estimated adjustment maps.

The fourth strand concerns sampling frequency. Chambers (2011) shows that cointegrating relationships and discrete-time representations can depend on the observation frequency. This point is especially important in fragmented high-frequency data, where one-second, five-second, and one-minute systems need not be simple rescalings of one another. The frequency-defect statistic $F_{k\delta} - F_{\delta}^{k}$ used below is a direct finite-sample diagnostic for this issue. It complements the existing sampling-frequency literature by asking whether estimated companion transitions behave as if they came from a common underlying transition mechanism.

Finally, the commutator diagnostic connects the price-discovery literature to an operator view of adjustment. Standard VECM analysis studies the common trend, the cointegration space, and the adjustment matrix. The commutator instead asks whether two market-specific update maps are order-invariant. In fragmented markets this is an economically natural question: if venue i's adjustment changes the state to which venue j reacts, then applying the two adjustment channels in opposite orders can produce different spread effects. This order-dependence is not summarized by Hasbrouck shares, realized-variance estimators, or conventional lead–lag correlations.

3 Data and Preprocessing

The data are synchronized one-second refreshed IBM prices from the multi-venue environment of Dias and Schweikert (2026), who study integrated variance estimation under fragmentation noise. The venue identifiers follow their notation: N is NYSE, T is Nasdaq, P is Arca, K is Cboe EDGX, and Z is Cboe BZX. The March 2021 sample contains 517,523 raw rows across 23 trading days, from 2021-03-01 10:45:00 to 2021-03-31 18:00:00. Table 1 reports the basic data diagnostics. The sample has no duplicate timestamps, no non-positive prices, no missing one-second timestamps within trading days, and no obvious return-jump flags. The only reported quality issue is 45 missing price cells in the five-price panel.

Log prices and within-day log returns are computed at the original 1-second grid and at resampled 5, 10, 30, and 60 second grids. Overnight and weekend gaps are treated as market-closure breaks, not as long returns, and are excluded from return and VECM lag observations. This choice keeps the return object tied to intraday price discovery: retaining calendar-time gaps would mix trading-time adjustment with market-closure intervals. Figure 1 plots the one-second price panel against a continuous trading-time observation index, so closed-market periods do not create visually misleading horizontal intervals. The figure shows the high-frequency comovement expected from five observations of the same asset, while the small separations between the lines show that venue-level discrepancies persist on the synchronized grid.

Table 1: Data diagnostics for the March 2021 panel

The table reports the integrity checks that define the usable synchronized panel. The absence of missing one-second timestamps, duplicate timestamps, non-positive prices, and obvious return-jump flags implies that the empirical analysis is not driven by basic timestamp or price-quality failures; the 45 missing price cells are the only reported data-quality issue.

Diagnostic	Value
Raw rows	517,523
Trading days	23
Missing one-second timestamps	0
Duplicate timestamps	0
Missing price cells	45
Non-positive prices	0
Obvious return-jump flags	0

Figure 1: Synchronized market-specific IBM prices at the one-second grid, plotted against a continuous trading-time observation index.

The horizontal axis is a trading-time observation index rather than calendar time. The five venue-specific prices move together closely, as expected for the same asset, but the small separations between the lines show persistent high-frequency venue dislocations.

4 Cointegration and VECM Benchmark

The benchmark targets the long-run restriction that five venue prices share one efficient-price component while temporary venue dislocations remain stationary. Following the cointegrated price-discovery literature, the log-price vector is modeled at each sampling frequency as a VECM (Johansen, 1991; Hasbrouck, 1995; Lütkepohl, 2005),

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{\ell=1}^{p-1} \Gamma_\ell \Delta X_{t-\ell} + c + \varepsilon_t,$$

where $\beta' X_t$ are stationary price-dislocation spreads, α contains error-correction loadings, and Γ_ℓ contains short-run return dynamics. Johansen trace tests estimate the cointegration rank. Table 2 shows that the selected rank is 4 at every sampling frequency, implying one common permanent component across the five markets. Lag length is selected by AIC, BIC, and HQ over candidate level lags $p = 1, \ldots, 5$. All three criteria select $p = 5$ at each frequency, with usable VECM sample sizes ranging from 517,399 observations at 1 second to 8,526 observations at 60 seconds.

The companion-form transition matrix F_δ is extracted from each estimated VECM. This matrix

is the one-step prediction map for the stacked state that contains the current price vector and the lagged price vectors needed by the VECM. It is useful because it puts long-run error correction and short-run lag dynamics into one linear object: applying F_δ once gives the fitted one-period evolution of the whole state at sampling interval δ. Figure 2 displays the first estimated one-second cointegration spread. Its fluctuations are small relative to the price level and remain centered around a stable dislocation band, consistent with the rank-four cointegration result. The figure also illustrates why the analysis is conducted in trading time: the economically relevant variation occurs over active trading observations rather than over overnight or weekend intervals.

Table 2: Johansen rank and VECM lag selection

The table summarizes the long-run rank and lag-length choice at each sampling frequency. Rank four means that the five venue prices share one permanent component, while the repeated choice $p = 5$ indicates that short-run lag dynamics remain relevant across all grids.

Frequency (s)	Selected rank	VECM lag p	Observations	BIC	HQ
1	4	5	517,399	-95.5625	-95.5645
5	4	5	103,401	-90.5130	-90.5210
10	4	5	51,651	-89.2701	-89.2902
30	4	5	17,151	-87.9345	-87.9723
60	4	5	8,526	-87.1327	-87.2008

Figure 2: First estimated cointegration spread at the one-second frequency, plotted against a continuous trading-time observation index.

The spread remains in a narrow stationary band instead of drifting with the price level. This visual pattern is consistent with the Johansen rank result and with the interpretation of $\beta' X_t$ as a temporary venue-level price dislocation.

5 Frequency Consistency

The frequency-consistency exercise asks whether the fitted transitions at different sampling grids contain the same predictive information. The benchmark is deliberately simple. Suppose the one-second and five-second systems were just two observations of the same stable linear adjustment mechanism. Then applying the one-second transition five times should give approximately the same state update as estimating the transition directly on five-second observations. More generally, if the systems were generated by a stable single linear transition mechanism, a k-step coarse-frequency transition should approximately satisfy (Chambers, 2011)

$$F_{k\Delta} \approx F_{\Delta}^{k}.$$

For each feasible pair of frequencies, define the frequency-defect matrix

$$D_{\Delta,k} = F_{k\Delta} - F_{\Delta}^{k}.$$

The matrix $D_{\Delta,k}$ is therefore a direct discrepancy measure. It is zero under exact frequency consistency and large when the coarser-grid transition contains dynamics that cannot be recovered by repeatedly applying the finer-grid transition. The Frobenius norm summarizes the total element-by-element discrepancy in the companion transition, while the spectral norm reports the largest discrepancy in any linear direction of the state. These are diagnostic norms, not structural parameters: they test whether the fitted discrete-time representations scale cleanly across grids. Block-bootstrap confidence intervals are computed using 99 replications and a block length of 300 observations. Table 3 reports the main defect estimates. The Frobenius norms range from 4.667 for the 5-to-30 second comparison to 5.866 for the 30-to-60 second comparison; spectral norms reach 2.672 for the 1-to-60 second comparison. To put these values on scale, the Frobenius defects are about 1.01 to 1.25 times the Frobenius norm of the corresponding coarse-grid transition matrix, and the spectral defects are about 1.11 to 1.68 times the corresponding spectral norm. The bootstrap intervals are tight and far from zero, so the discrepancies are not well interpreted as numerical noise around a frequency-invariant transition operator.

Figure 3 visualizes the same result across all compatible frequency pairs. The heatmap has no near-zero region: defects remain large whether the comparison starts from 1 second, 5 seconds, 10 seconds, or 30 seconds. The fitted transitions therefore violate the stable single-generator bench-

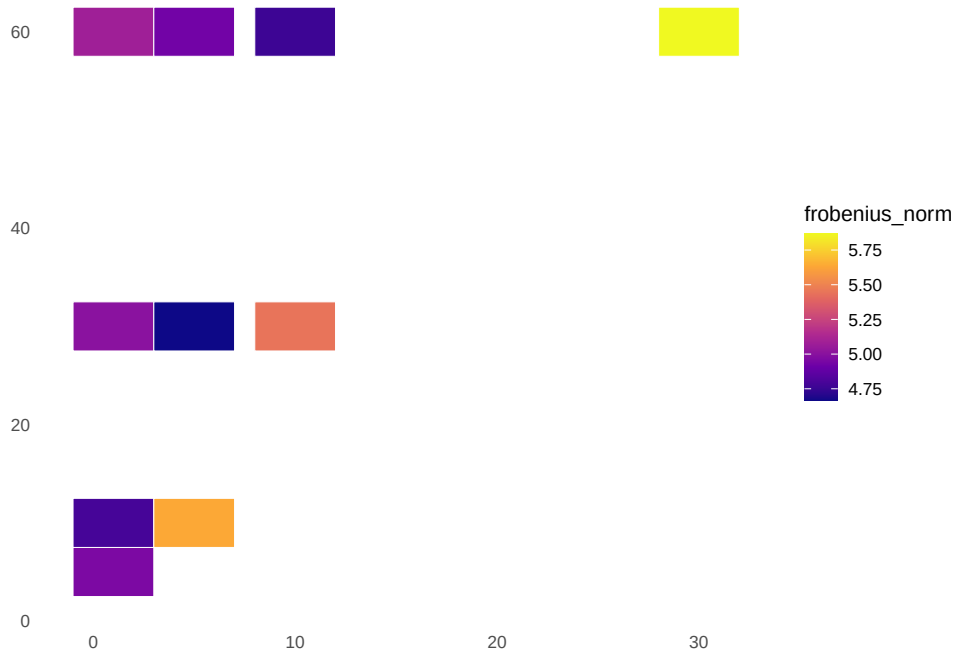
Table 3: Frequency-defect diagnostics

The table compares estimated coarse-grid transitions with powers of finer-grid transitions. Large Frobenius and spectral norms, together with bootstrap intervals far from zero, indicate that the same linear transition mechanism does not scale cleanly across sampling frequencies.

Base (s)	Target (s)	k	$\|D\|_F$	$\|D\|_2$	95% CI for $\|D\|_F$
1	5	5	4.960	2.278	[4.948, 4.974]
1	10	10	4.791	2.419	[4.775, 4.821]
1	30	30	5.003	2.651	[4.979, 5.070]
1	60	60	5.079	2.672	[5.059, 5.323]
5	10	2	5.631	1.865	[5.628, 5.641]
30	60	2	5.866	1.952	[5.864, 6.109]

Figure 3: Frequency-defect Frobenius norms across compatible sampling-frequency pairs.

Each cell reports the size of $F_{k\delta} - F_{\delta}^{k}$ for a compatible pair of sampling intervals. Darker cells indicate larger deviations from the single-generator benchmark; the absence of a near-zero region shows that frequency dependence is pervasive.



mark. In plain terms, the model estimated at a coarser grid cannot be recovered simply by applying the finer-grid model repeatedly.

6 Lead–Lag Correlations and Hasbrouck Shares

The next step separates two objects that are often described with the same price-discovery language. Pairwise lead–lag return correlations are computed at the one-second grid across short lags. These correlations measure timing rather than permanent-innovation contributions (Hasbrouck, 1995; Dias et al., 2021). For $h > 0$, define

$$\rho_{i \to j}(h) = \mathrm{corr}(r_{i,t}, r_{j,t+h}).$$

This statistic is a short-horizon timing measure, not a structural venue-contribution share. To avoid overinterpreting a fragile aggregate rank, Table 4 reports unordered-pair asymmetries: the peak one-second directional edge and the integrated signed advantage over the short-lag window. The largest peak edge is $P \to N$ at 0.110, followed by $T \to N$ at 0.069, $P \to Z$ at 0.059, $K \to N$ at 0.052, and $Z \to N$ at 0.045. These estimates indicate systematic one-second timing asymmetries, especially for pairs involving N, but they are return-correlation asymmetries on the synchronized grid rather than permanent-price contribution shares.

Hasbrouck information shares are computed separately from the permanent-price representation of the VECM. Because the innovation covariance matrix is not diagonal, Table 5 reports lower and upper information-share bounds across Cholesky orderings. At the one-second frequency, N has the largest midpoint share, 0.474, with a wide interval from 0.088 to 0.860. T follows with midpoint 0.351, Z with 0.297, K with 0.254, and P with 0.238. This ordering is consistent with the venue-contribution pattern emphasized by Dias and Schweikert (2026), while the wide bounds show that Hasbrouck allocations depend strongly on contemporaneous innovation ordering.

Figure 4 plots the full lead–lag correlation profiles and shows that the largest directional edges occur at the shortest positive lag. Figure 5 extends the Hasbrouck comparison across sampling frequencies. As the sampling interval increases, midpoint shares become more compressed even though N remains at the top of the ordering. The two figures therefore separate two notions of price discovery. Hasbrouck shares summarize permanent-innovation contributions, whereas lead–lag correlations summarize short-run timing asymmetries.

Table 4: One-second pairwise lead–lag asymmetry

The table reports directional timing asymmetries, not structural price-discovery shares. A positive peak direction means that one market's return is most correlated with another market's subsequent return at the one-second lag; the integrated edge aggregates this signed advantage over the short-lag window.

Pair	Peak direction	Peak edge	Integrated edge
N/T	T → N	0.069	-0.133
N/K	K → N	0.052	-0.042
N/P	P → N	0.110	-0.172
N/Z	Z → N	0.045	-0.088
T/K	T → K	0.005	0.052
T/P	P → T	0.042	-0.054
T/Z	T → Z	0.019	0.035
K/P	P → K	0.040	-0.088
K/Z	K → Z	0.014	-0.015
P/Z	P → Z	0.059	0.082

Table 5: One-second Hasbrouck information-share bounds

The table reports lower and upper Hasbrouck bounds across Cholesky orderings and uses the midpoint only as a compact summary. The wide intervals show that contemporaneous innovation ordering matters, even though N has the highest midpoint in this sample.

Market	Rank	Lower	Upper	Midpoint
N	1	0.088	0.860	0.474
T	2	0.035	0.668	0.351
Z	3	0.013	0.581	0.297
K	4	0.013	0.495	0.254
P	5	0.026	0.449	0.238

Figure 4: Pairwise one-second lead–lag return correlations.

The curves compare return correlations after shifting one venue relative to another. Peaks at the shortest positive lag reveal timing asymmetries on the refreshed grid, but these correlations should not be interpreted as permanent-price contribution shares.

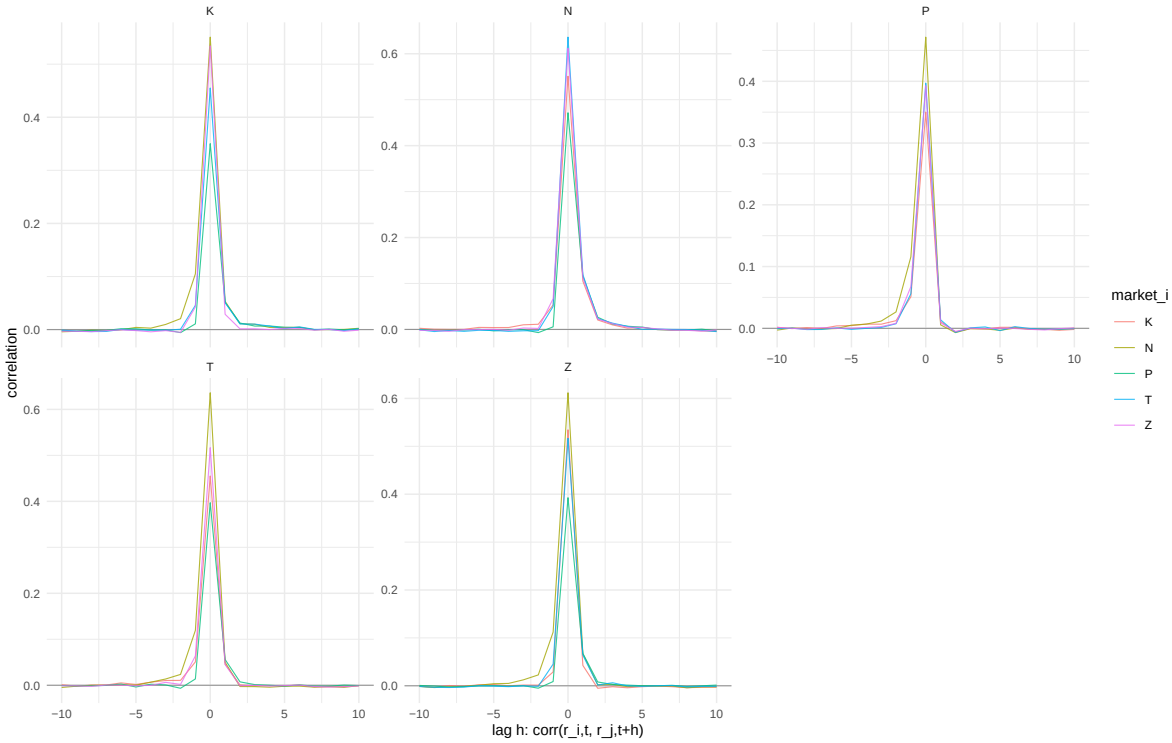
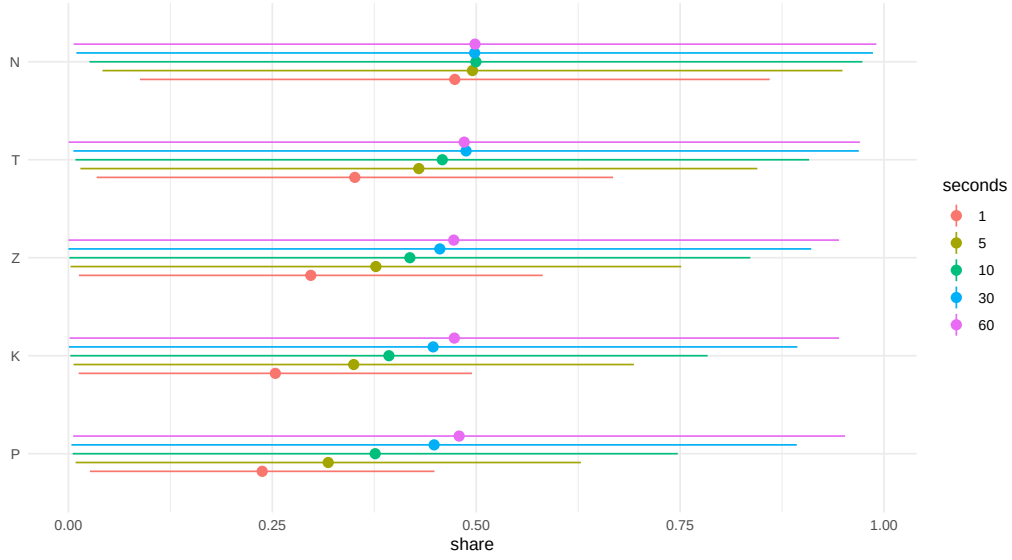


Figure 5: Hasbrouck information-share bounds by market and sampling frequency.

The intervals are Cholesky-ordering bounds, with midpoints used only as a compact summary. N remains the highest-midpoint venue, while the wide bounds show that contemporaneous innovation ordering remains important for allocating permanent-price variation.



7 Theory of the Commutator Diagnostic

7.1 Motivation

The starting point is the fragmentation-noise mechanism of Dias and Schweikert (2026). In their framework, prices for the same asset are observed in several venues, share a latent efficient price, and correct deviations through a cointegrated drift. This drift is not just a nuisance term: it is the mechanism through which venue-level prices learn about and revert toward the common efficient price. A standard VECM summarizes this learning through a single linear transition operator. The commutator diagnostic asks a different question: does the estimated transition imply that market-specific adjustment channels can be applied in any order, or does the order of market updates matter?

In linear algebra, two operators commute when $AB = BA$. Commutation means that applying A then B gives the same result as applying B then A. Non-commutation is therefore a direct mathematical representation of order-dependence. In the present setting, the relevant operators are not observed event-time trading operators. They are constructed from the fitted VECM transition matrix as empirical market-update operators. The resulting commutator is a plug-in diagnostic of order-dependence in the estimated fragmented price system, not a causal estimate of trade sequencing across venues. Its value is that it converts a qualitative feature of fragmented trading–the possibility that adjustment channels interact asymmetrically–into an estimable matrix object with interpretable projections.

7.2 From the VECM to a State-Space Operator

At sampling interval δ, write the fitted VECM with $m = 5$ markets and level lag p as

$$\Delta X_t = \alpha_\delta \beta_\delta' X_{t-1} + \sum_{\ell=1}^{p-1} \Gamma_{\delta,\ell} \Delta X_{t-\ell} + c_\delta + \varepsilon_{\delta,t}.$$

For the operator calculation, the intercept is omitted and the model is written in companion form on the augmented state

$$\mathcal{X}_t = \begin{pmatrix} X_t' & X_{t-1}' & \cdots & X_{t-p+1}' \end{pmatrix}' \in \mathbb{R}^{mp}, \qquad \mathcal{X}_t = F_\delta \mathcal{X}_{t-1} + \eta_{\delta,t}.$$

The top m rows of F_δ determine how the five current market prices update as functions of the lagged state; the remaining rows shift lags mechanically. This separation identifies market-specific update rows without imposing a recursive ordering among venues. The construction below uses only these fitted rows. It does not assume that actual trades arrive one venue at a time in the order used by the diagnostic.

Let $d = mp$, let e_i be the i-th canonical basis vector in $\mathbb{R}^d$, and let $E_i = e_i e_i'$ for $i = 1, \ldots, m$. Define $A_\delta = F_\delta - I_d$. The market-$i$ row-update operator is

$$U_{\delta,i} = I_d + E_i A_\delta.$$

Applied to a state vector, $U_{\delta,i}$ leaves every coordinate unchanged except the current-price coordinate for market i, which is updated using the same row of F_δ that the full companion system would use. Equivalently, if $a_{\delta,i}' = e_i' A_\delta$ is the fitted adjustment row for market i, then

$$U_{\delta,i} x = x + e_i a_{\delta,i}' x.$$

This construction yields a local counterfactual object: update market i using its estimated adjustment rule while holding the other current-market coordinates and lag coordinates fixed. It is local because it isolates one fitted row of the linear transition. It is counterfactual because the full VECM updates all rows at once; the diagnostic asks what the fitted system implies if the row updates are separated.

7.3 Commutators as Order-Dependence

For markets i and j, define

$$C_{\delta,ij} = U_{\delta,j} U_{\delta,i} - U_{\delta,i} U_{\delta,j}.$$

If $C_{\delta,ij} = 0$, then the fitted market-i and market-j updates commute: the local transition implied by updating i then j is identical to updating j then i. If $C_{\delta,ij} \neq 0$, the fitted transition is order-dependent. Because $U_{\delta,i} = I_d + E_i A_\delta$, the first-order terms cancel and

$$C_{\delta,ij} = E_j A_\delta E_i A_\delta - E_i A_\delta E_j A_\delta.$$

This cancellation is why the diagnostic is not just another measure of how large the two individual adjustment rows are. Acting on a state x, the commutator can be written as

$$C_{\delta,ij}x = e_j(e_j' A_\delta e_i)(a_{\delta,i}' x) - e_i(e_i' A_\delta e_j)(a_{\delta,j}' x).$$

The first term is the effect of market j's fitted update reacting to the part of the state changed by market i; the second term is the reverse channel. Thus the commutator captures a second-order interaction between market-specific adjustment rows. A large commutator can arise when these two channels have different strength or sign. This is the kind of asymmetric fragmentation mechanism that a single Hasbrouck information-share ranking may not reveal. The object is also distinct from a lead–lag correlation: lead–lag correlations compare observed returns across time shifts, whereas $C_{\delta,ij}$ compares the estimated adjustment maps themselves.

The formal target is the row-update commutator of the linear transition operator at a chosen sampling interval, together with its projected distribution over observed trading-time states. This is an estimand for the fitted discrete-time adjustment system, not for the causal sequence of event-time trades. For fixed δ, p, and cointegration space, the object is

$$\mathcal{D}_{\delta,ij} = \{C_{\delta,ij},\, B_\delta C_{\delta,ij}\mathcal{X}_t,\, Q_\delta C_{\delta,ij}\mathcal{X}_t\}_t\,.$$

Here B_δ and Q_δ, defined below, project the commutator effect onto stationary spreads and the permanent-price direction, respectively. The set ranges over usable within-day companion states. It records three pieces of information: the matrix that measures non-commutation, the time series of implied spread effects, and the time series of implied permanent-price effects. It measures how much the two row-update orderings differ, and whether that difference loads onto economically interpretable directions. The empirical estimator is the plug-in matrix

$$\widehat{C}_{\delta,ij} = \widehat{U}_{\delta,j}\widehat{U}_{\delta,i} - \widehat{U}_{\delta,i}\widehat{U}_{\delta,j}, \qquad \widehat{U}_{\delta,i} = I_d + E_i(\widehat{F}_\delta - I_d).$$

The matrix norms $\|\widehat{C}_{\delta,ij}\|_F$ and $\|\widehat{C}_{\delta,ij}\|_2$ summarize pairwise order-dependence in the estimated transition system. These norms rank pair strength. The economically relevant question is where the order-dependent component loads: stationary price dislocations or the permanent-price direction.

7.4 Spread and Permanent-Price Projections

Matrix norms show that two update operators fail to commute, but they do not say whether the non-commutation matters for economically meaningful directions. We therefore project the commutator onto the cointegration-spread space and onto the permanent-price direction. Embed the estimated spread matrix in the augmented state as

$$B_\delta = \begin{pmatrix} \widehat{\beta}_\delta' & 0 & \cdots & 0 \end{pmatrix}.$$

For an observed augmented state $\mathcal{X}_t$, the spread-projected commutator effect is

$$G^{\beta}_{\delta,ij,t} = B_\delta \widehat{C}_{\delta,ij} \mathcal{X}_t.$$

When $p = 1$, this reduces to the simpler expression $\widehat{\beta}_\delta' \widehat{C}_{\delta,ij} X_t$. With $p > 1$, the projection correctly accounts for lagged-state dependence in the companion system. Analogously, if q_δ denotes the normalized permanent-price direction implied by the orthogonal complement of $\widehat{\beta}_\delta$, define

$$Q_\delta = \begin{pmatrix} q_\delta' & 0 & \cdots & 0 \end{pmatrix}, \qquad G^{q}_{\delta,ij,t} = Q_\delta \widehat{C}_{\delta,ij} \mathcal{X}_t.$$

The empirical distribution of $G^{\beta}_{\delta,ij,t}$ is the main object for studying whether order-dependence generates spread dislocations, asymmetric higher moments, or tail risk. The reported statistics include variance, skewness, excess kurtosis, quantiles, and tail probabilities for each pair and spread projection. Block-bootstrap intervals are computed by resampling contiguous observations, re-estimating the VECM, reconstructing $\widehat{F}_\delta$, and recomputing the projected commutator statistics. This preserves local dependence while propagating estimation uncertainty through the entire commutator construction.

8 Empirical Commutator Results

Table 6 reports the five largest one-second commutator pairs by beta-projected variance. The N/Z pair has the largest projected variance, 1.215×10^{-11}, followed by P/Z at 8.732×10^{-12}, T/Z at 7.578×10^{-12}, and N/T at 6.795×10^{-12}. Taking the square root of the N/Z variance gives 3.49×10^{-6} log points, or about 0.035 basis points, on the combined spread-projection scale. Across the individual N/Z spread projections, the largest standard deviation is 0.019 basis points and the

largest absolute 1% tail quantile is 0.080 basis points. The diagnostic therefore identifies systematic order-dependence in fitted adjustment, not a large mechanical arbitrage signal. Norm-based and projection-based rankings are related but not identical: N/T has the largest Frobenius norm, 0.0358, whereas N/Z has the largest beta-projected variance. This difference matters because the projection asks whether non-commutation loads onto economically relevant spread directions, not only whether the full companion-state matrices differ.

Table 6: Pairwise commutator strength at the one-second frequency

The table ranks the largest commutator pairs by beta-projected variance. Comparing matrix norms with projected variance shows that the strongest full-state non-commutation need not be the strongest economically relevant spread dislocation.

Pair	$\|C\|_F$	$\|C\|_2$	Beta-projected variance
N/Z	0.0244	0.0217	1.215×10^{-11}
P/Z	0.0168	0.0130	8.732×10^{-12}
T/Z	0.0242	0.0207	7.578×10^{-12}
N/T	0.0358	0.0318	6.795×10^{-12}
N/P	0.0219	0.0184	4.099×10^{-12}

Figure 6 plots the full pairwise commutator strength matrix. The heatmap shows that order-dependence is not concentrated in a single idiosyncratic pair: N/T, N/Z, T/Z, and T/P all display visible commutator strength, while K/Z is comparatively weak. Projected commutator effects are computed along spread directions and the permanent-price direction. In the companion-state notation of the theory section, the spread-projected effect is $G^{\beta}_{\delta,ij,t} = B_{\delta} C_{\delta,ij} \mathcal{X}_t$. Suppressing lags only for exposition, this corresponds to

$$G^{\beta}_{ij,t} = \beta' C_{ij} X_t.$$

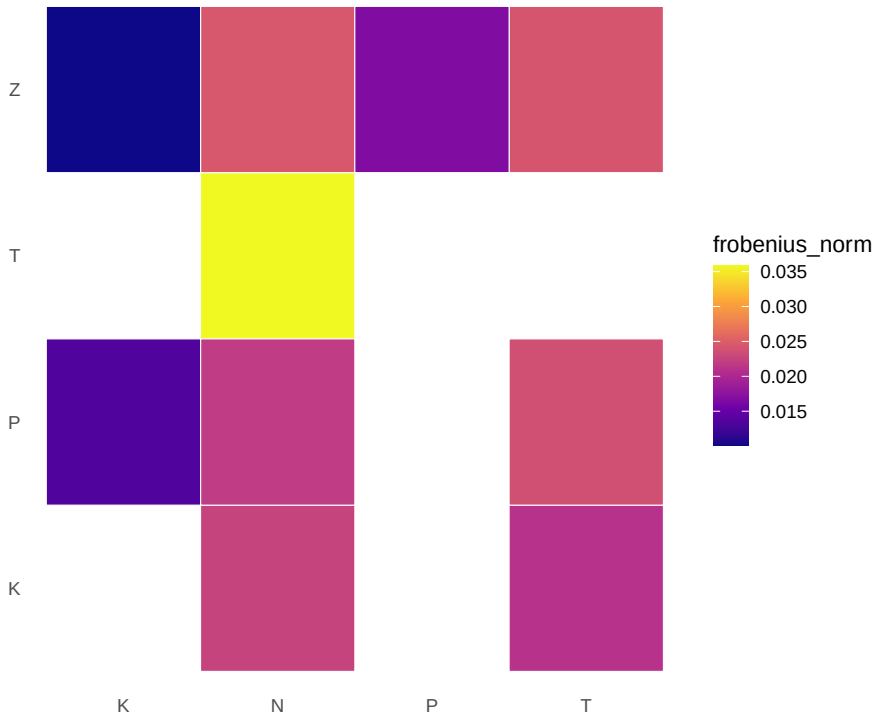
Across spread projections, the largest absolute skewness is 0.483, the largest excess kurtosis is 15.184, and the largest empirical three-standard-deviation tail probability is 0.018. These higher-moment effects show that order-dependence is not only a variance phenomenon; it also appears in asymmetric and heavy-tailed spread dislocations.

9 Volatility and Higher Moments

Gaussian innovations are used only as a benchmark. VECM residual diagnostics reject normality for every market and frequency. This is consistent with the high-frequency volatility literature,

Figure 6: Pairwise commutator strength matrix.

Off-diagonal cells summarize the strength of pairwise non-commutation between market-update operators. The visible mass outside a single pair indicates that order-dependence is a system-wide feature of the fitted fragmented price dynamics.



which treats microstructure noise, jumps, and conditional heteroskedasticity as central empirical features rather than minor deviations from a Gaussian benchmark (Hansen and Lunde, 2006; Bandi and Russell, 2008; Barndorff-Nielsen et al., 2008; Christensen et al., 2010). Ljung–Box tests on squared residuals and ARCH-style diagnostics also indicate volatility clustering throughout the panel. Table 7 reports the residual higher-moment summary. Median residual excess kurtosis declines with coarser sampling but remains positive: 11.39 at 1 second, 4.18 at 5 seconds, 3.57 at 10 seconds, 2.33 at 30 seconds, and 2.37 at 60 seconds. The volatility-clustering share is 1.00 at every frequency, so the Gaussian VECM benchmark leaves substantial conditional heteroskedasticity in the residuals.

Table 7: Residual higher moments and volatility clustering

The table summarizes departures from Gaussian homoskedastic residual behavior. Excess kurtosis remains positive at all frequencies, and the volatility-clustering share of 1.00 indicates that conditional heteroskedasticity survives the linear cointegrated benchmark.

Frequency (s)	Median residual excess kurtosis	Share with volatility clustering
1	11.39	1.00
5	4.18	1.00
10	3.57	1.00
30	2.33	1.00
60	2.37	1.00

Figure 7: Mean daily realized variance by sampling frequency and market.

The figure compares mean daily realized variance across venues and sampling grids. Realized variance is highest at the finest grid for several venues, especially P, which is consistent with venue-specific high-frequency noise rather than a single common volatility profile.

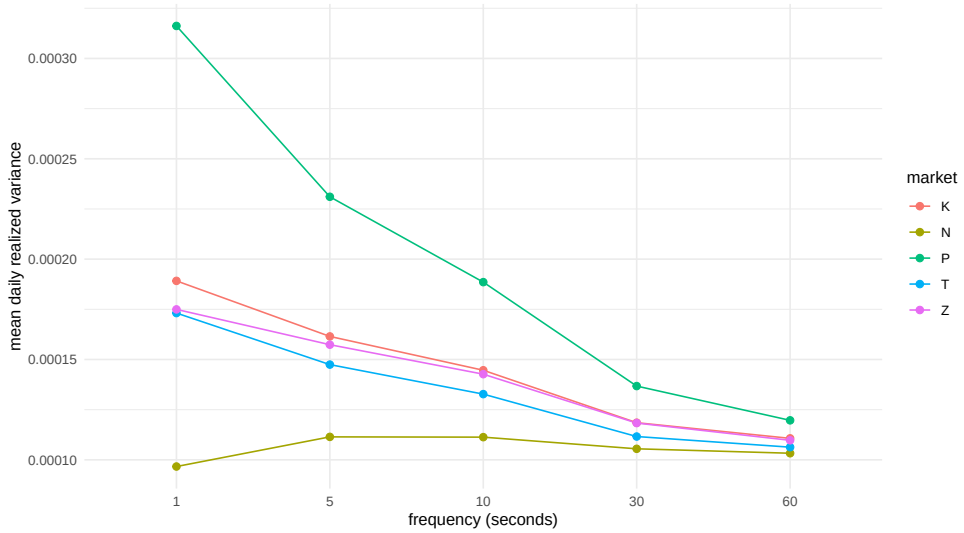


Figure 7 shows that realized variance is materially frequency dependent and differs across venues.

Averaging across markets and days, mean daily realized variance is 1.90×10^{-4} at 1 second and falls to 1.10×10^{-4} at 60 seconds. At the one-second frequency, P has the largest mean realized variance, 3.16×10^{-4}, while N has the smallest, 9.67×10^{-5}. The pattern is consistent with stronger high-frequency noise in some venue-specific series rather than a single common volatility curve shared by all markets.

Figure 8: Efficient-price realized variance by sampling frequency.
The common-component realized variance is much more stable across frequencies than the venue-specific realized variances. This stability is consistent with the estimated one-dimensional permanent component, although the level still varies with sampling interval.

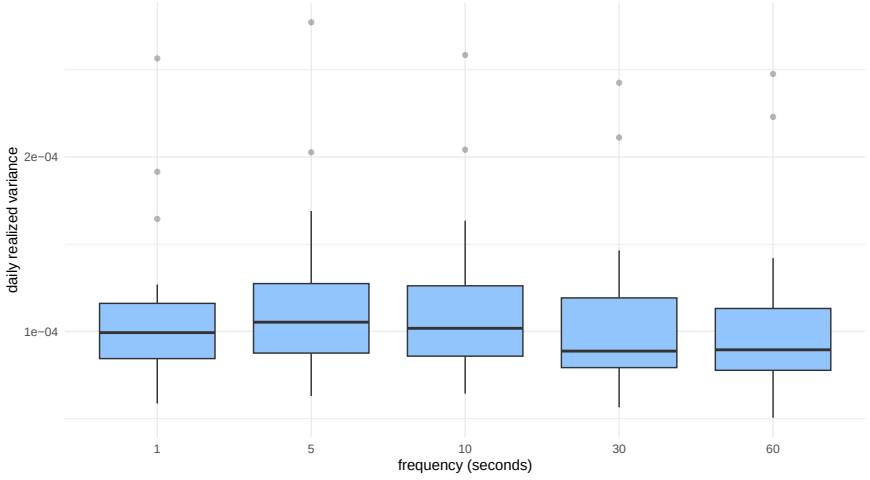


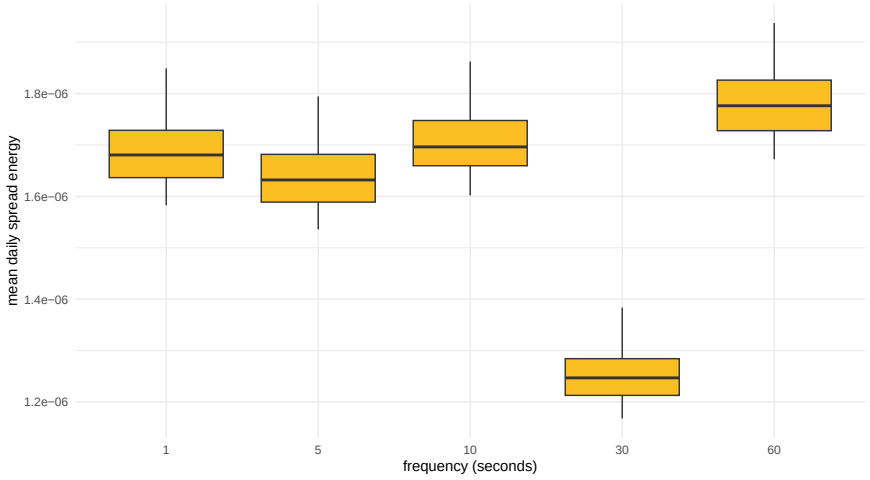
Figure 8 reports the efficient-price realized variance implied by the common component. Its daily mean is 1.067×10^{-4} at 1 second, 1.142×10^{-4} at 5 seconds, 1.131×10^{-4} at 10 seconds, 1.057×10^{-4} at 30 seconds, and 1.032×10^{-4} at 60 seconds. Figure 9 reports the spread/fragmentation energy. Mean spread energy is 1.690×10^{-6} at 1 second, 1.642×10^{-6} at 5 seconds, 1.711×10^{-6} at 10 seconds, 1.255×10^{-6} at 30 seconds, and 1.783×10^{-6} at 60 seconds. These diagnostics indicate that the cointegrated VECM captures first-order long-run restrictions, but leaves economically relevant non-Gaussian and time-varying volatility structure in the residuals. The frequency and commutator results further imply that fragmentation volatility cannot be summarized only by permanent-price innovation shares.

10 Conclusion

The empirical message is straightforward. The five IBM venue prices move around one common price, so the standard cointegrated price-discovery benchmark is a useful starting point. But the

Figure 9: Spread/fragmentation energy by sampling frequency.

The figure summarizes the average size of stationary venue dislocations. The non-monotone frequency pattern shows that fragmentation energy is not eliminated by moving to a coarser grid and should be treated separately from permanent-price volatility.



short-run adjustment process is not stable across sampling grids. A five-second or one-minute model does not look like a simple accumulation of one-second adjustments, and the bootstrap results show that this gap is too large to dismiss as estimation noise. Sampling frequency is therefore not a harmless technical choice in this setting; it changes the adjustment pattern that the researcher estimates.

The results also show that a single price-discovery ranking is incomplete. Hasbrouck shares summarize which venues contribute most to the common permanent price, and they remain useful for that purpose. They do not show whether one venue tends to move before another in the short run, or whether the fitted adjustment after a shock depends on which venue reacts first. The commutator diagnostic is meant to capture this second issue in a simple way: for some venue pairs, applying the two estimated adjustment channels in one order gives a different temporary price dislocation than applying them in the opposite order. Those order effects are visible in spread variation and in the tails of the fitted dislocation measures.

This is the main difference from Dias and Schweikert (2026). Their paper treats fragmentation as a problem for volatility measurement and uses the multi-venue error-correction system to recover a cleaner efficient-price series before estimating integrated variance. The present paper studies the error-correction system itself. It asks whether the adjustment mechanism is stable across sampling frequencies and whether venue-specific adjustment channels interact in an order-dependent way.

The contribution is therefore not another integrated-variance estimator, nor a competing estimate of the same efficient-price volatility. It is evidence that the mechanism used to filter fragmentation noise also contains economically interpretable price-discovery information that is lost when the system is reduced to a permanent-price share or a defragmented volatility input.

The practical implication is that empirical price-discovery and volatility analysis should treat the sampling grid, market-closure handling, and single-ranking reductions as substantive choices. Intraday returns should be constructed in trading time, with overnight and weekend closures excluded from the return object unless the research question explicitly concerns close-to-open price changes. Hasbrouck shares remain useful for summarizing permanent-price contributions, but they should be reported together with short-run timing measures and frequency-sensitivity checks. For risk measurement, the evidence implies that venue-specific high-frequency variation and spread energy can persist even when a common efficient-price component is well defined.

The commutator diagnostic adds a practical layer. In plain terms, it identifies pairs of venues for which the answer changes when we ask which venue reacts first. These cases are useful because they point to market interactions that a single ranking would hide. In applications, this can guide robustness checks for volatility estimates, inform market-quality comparisons across venues, and flag pairs where temporary price dislocations deserve closer attention.

The same information may also have trading value, but only under stringent execution conditions. The timing asymmetries and order-dependent spread dislocations would matter for trading only if they predict price movements that are larger than bid–ask spreads, fees and rebates, latency, queue-position costs, adverse selection, and inventory risk. The present evidence is therefore not a trading-profit claim. It instead defines a natural out-of-sample test: use the lead–lag and order-dependence signals to predict which venue is temporarily lagging, trade only when the predicted dislocation exceeds a conservative cost threshold, and close the position once the spread reverts or after a fixed short horizon. Such a test should use strict chronology, no look-ahead estimation, executable quotes rather than refreshed prices, and conservative transaction-cost assumptions.

The main limitation is interpretive. The diagnostic is estimated from a statistical model of synchronized prices. It does not prove that one venue causally moves before another in event time. A natural extension is therefore to estimate comparable diagnostics directly in event time or trade time and compare them with the synchronized-grid results reported here.

The evidence is deliberately narrow. It establishes frequency-dependent and order-dependent adjustment in a synchronized multi-venue IBM setting, with a bounded February screen as a month check. The contribution is therefore not a market-wide prevalence estimate. It is a reproducible framework for testing whether permanent-price shares, sampling-frequency effects, and order-dependent adjustment remain distinct in broader multi-asset panels.

A Lag-Sensitivity Screen

The main analysis uses the information-criterion choice $p = 5$ at every sampling grid. To check whether the frequency-defect and commutator conclusions are tied mechanically to that lag choice, Table 8 re-estimates the March 2021 one- and five-second VECMs for common lag choices $p = 1, \ldots, 5$, using the same cointegration spaces as the main estimates. The screen reports the 1-to-5 second transition defect and the strongest one-second beta-projected commutator pair at each lag.

Table 8: Lag-sensitivity screen for the 1-to-5 second comparison

The table is a bounded sensitivity screen, not a replacement for the selected-lag bootstrap analysis. The first two columns report norms of $F_{5\delta} - F_{\delta}^{5}$. The final columns report the pair with the largest beta-projected commutator variance at the one-second grid under the same lag choice.

p	n_{1s}	n_{5s}	$\|D\|_F$	Top pair	Beta variance
1	517,491	103,493	0.205	T/Z	3.109×10^{-12}
2	517,468	103,470	1.243	N/Z	1.147×10^{-11}
3	517,445	103,447	2.945	N/Z	1.195×10^{-11}
4	517,422	103,424	3.981	N/Z	1.207×10^{-11}
5	517,399	103,401	4.960	N/Z	1.215×10^{-11}

The lag screen shows that magnitudes are lag-sensitive, as expected in a high-frequency VECM with persistent short-run adjustment. The underfit $p = 1$ specification produces a smaller 1-to-5 second defect and a different top beta-projected commutator pair. From $p = 2$ onward, however, the strongest beta-projected pair is consistently N/Z, and the largest Frobenius-norm commutator pair remains N/T in the generated table. Thus the selected-lag commutator result is not driven only by the final increment from $p = 4$ to $p = 5$, although the defect magnitude should be interpreted jointly with the lag-selection evidence in Table 2.

B Month-Robustness Screen

The main estimates use March 2021 because this is the month used for the full multi-frequency and bootstrap analysis. To check whether the core long-run and frequency-dependence results are

specific to that calendar month, Table 9 reports a bounded screen that re-estimates the rank and 1-to-5 second transition defect for February and March 2021. The screen is deliberately cheaper than the main analysis: it uses the 1 and 5 second grids, caps the VECM level lag at $p = 2$, and does not compute bootstrap intervals.

Table 9: Month-robustness screen at the 1 and 5 second grids

The screen uses the same synchronized five-market IBM panel construction as the main analysis. The final column reports the Frobenius norm of the 1-to-5 second defect $F_{5\delta} - F_{\delta}^{5}$ under the common lag cap $p = 2$. The exercise is an external-validity screen, not a replacement for the full March 2021 bootstrap analysis.

Month	Days	Observations 1s/5s	Rank 1s	Rank 5s	Defect
2021-02	19	427,504 / 85,512	4	4	1.147
2021-03	23	517,514 / 103,516	4	4	1.243

The screen preserves the two core qualitative facts. First, both months select cointegration rank four at both grids, leaving one common permanent component in the five-market system. Second, the 1-to-5 second transition defect remains economically non-trivial in both months. The additional month therefore supports the paper's central claim that the common permanent-price benchmark does not exhaust the information in fragmented price discovery, although a full external-validity exercise should extend the bootstrap analysis to more assets and months.

C Illustrative Monetizability Screen

This appendix reports a deliberately limited monetizability screen. The exercise uses the first 11 trading days of March 2021 to select simple rules and the remaining 12 trading days for evaluation. It uses refreshed synchronized prices, not executable bid and ask quotes. The results therefore measure whether the documented timing and dislocation patterns contain gross midquote predictability; they do not establish that the patterns are tradable after bid–ask spreads, exchange fees and rebates, latency, queue position, adverse selection, inventory constraints, or market-impact costs.

Two simple rule classes are considered. The lead–lag timing basket selects the five strongest first-half one-second directional correlations and holds the lagging venue for one second in the direction implied by the leading venue return, using only signals above the first-half 75th percentile of absolute leader returns. The spread-convergence basket selects the five first-half cross-venue spreads with the highest break-even transaction-cost capacity and trades one-second convergence only when the absolute spread exceeds its first-half 90th percentile. In both cases, the key statistic is the break-

even one-way cost in basis points: the all-in cost per unit of traded notional that would reduce the gross test-period profit to zero. Table 10 reports the out-of-sample screen.

Table 10: Illustrative monetizability screen on the second half of March 2021

The table reports gross midquote performance, not executable trading performance. A strategy is economically plausible only if its all-in one-way transaction cost is below the reported break-even cost. The lead–lag rule has very little cost capacity, while the spread-convergence rule has more room but still requires execution quality that cannot be verified with refreshed prices alone.

Strategy	Active	Turnover	Daily bps	t-stat	BE cost
Lead–lag timing basket	0.332	8,114.9	499.3	14.17	0.0615
Spread-convergence basket	0.256	1,993.5	1,488.9	9.95	0.7469

The screen suggests that the empirical patterns contain gross statistical predictability, but only under restrictive execution conditions. The lead–lag basket turns over more than 8,000 position units per day and has a break-even one-way cost of only 0.0615 basis points, which is unlikely to survive ordinary crossing costs. The spread-convergence basket has a larger break-even cost, 0.7469 basis points, and lower turnover, but even this figure is based on synchronized refreshed prices rather than executable quotes. A credible trading implementation would require quote-level backtesting with strict chronology, explicit maker/taker fees and rebates, latency assumptions, queue-position modeling, and inventory controls. The appropriate interpretation is that the results motivate a strategy design for further testing, not that they demonstrate an exploitable arbitrage.

Acknowledgements

We gratefully acknowledge access to TAQ provided by DALAHO, University of Hohenheim. We also thank Karsten Schweikert for his support in acquiring the data through DALAHO.

Data Availability and Replication

All main estimates use the March 2021 IBM multi-venue sample described above and the venue definitions in Dias and Schweikert (2026); the bounded month-robustness screen additionally uses February 2021. The underlying database extracts are subject to DALAHO access and use conditions and cannot be redistributed. The replication code is organized so authorized users can rebuild the tables and figures from a local raw-data archive; public replication materials should therefore exclude restricted raw and price-level data and include only code, documentation, and redistributable derived outputs. The reported bootstrap intervals use 99 block-bootstrap replications

with block length 300 and random seed 202103. The replication documentation gives the rebuild commands and maps the main claims to generated outputs.

Declaration of Generative AI and AI-Assisted Technologies in the Manuscript Preparation Process

During the preparation of this work, the author used OpenAI Codex to support manuscript editing, code and package checks, and documentation drafting. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the submission. No AI tool was used to generate, fabricate, or alter the underlying research data.

References

Bandi, F. M., and Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies*, 75(2), 339–369. doi:10.1111/j.1467-937X.2008.00474.x.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6), 1481–1536. doi:10.3982/ECTA6495.

Chambers, M. J. (2011). Cointegration and sampling frequency. *The Econometrics Journal*, 14(2), 156–185. doi:10.1111/j.1368-423X.2010.00329.x.

Christensen, K., Kinnebrock, S., and Podolskij, M. (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of Econometrics*, 159(1), 116–133. doi:10.1016/j.jeconom.2010.05.001.

Dias, G. F., and Schweikert, K. (2026). Integrated variance estimation for assets traded in multiple venues. *Journal of Econometrics*, 255, 106244. doi:10.1016/j.jeconom.2026.106244.

Dias, G. F., Fernandes, M., and Scherrer, C. M. (2021). Price discovery in a continuous-time setting. *Journal of Financial Econometrics*, 19(5), 985–1008. doi:10.1093/jjfinec/nbz030.

Hansen, P. R., and Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2), 127–161. doi:10.1198/073500106000000071.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *Journal of Finance*, 50(4), 1175–1199. doi:10.1111/j.1540-6261.1995.tb04054.x.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59(6), 1551–1580. doi:10.2307/2938278.

Menkveld, A. J. (2016). The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics*, 8, 1–24. doi:10.1146/annurev-financial-121415-033010.

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics*, 116(2), 257–270. doi:10.1016/j.jfineco.2015.01.003.

O'Hara, M., and Ye, M. (2011). Is market fragmentation harming market quality? *Journal of Financial Economics*, 100(3), 459–474. doi:10.1016/j.jfineco.2011.02.006.

Scherrer, C. M. (2021). Information processing on equity prices and exchange rate for cross-listed stocks. *Journal of Financial Markets*, 54, 100634. doi:10.1016/j.finmar.2021.100634.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer. doi:10.1007/978-3-540-27752-1.