School of Economics Working Paper 2024-01

An Econometric Analysis of Volatility Discovery

Gustavo Fruet Dias *
Fotis Papailias **
Cristina Scherrer ***

* University of East Anglia
** King's College London
*** London School of Economics

University of East Anglia

SCHOOL OF ECONOMICS

School of Economics
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/economics

An Econometric Analysis of Volatility Discovery

Gustavo Fruet Dias*
School of Economics – University of East Anglia (UEA)
Fotis Papailias
King's Business School – King's College London
Cristina Scherrer
Department of Finance – London School of Economics (LSE)

December 4, 2023

Abstract

We investigate information processing in the stochastic process driving stock's volatility (volatility discovery). We apply fractionally cointegration techniques to decompose the estimates of the market-specific integrated variances into an estimate of the common integrated variance of the efficient price and a transitory component. The market weights on the common integrated variance of the efficient price are the volatility discovery measures. We relate the volatility discovery measure to the price discovery framework and formally show their roles on the identification of the integrated variance of the efficient price. We establish the limiting distribution of the volatility discovery measures by resorting to both long span and in-fill asymptotics. The empirical application is in line with our theoretical results, as it reveals that trading venues incorporate new information into the stochastic volatility process in an individual manner and that the volatility discovery analysis identifies a distinct information process than that based on the price discovery analysis.

Keywords: long memory, fractionally cointegrated vector autoregressive model, realized measures, market microstructure, price discovery, high-frequency data, double asymptotics

Acknowledgements: The authors are grateful to the Editor Prof. Atsushi Inoue, an anonymous associate editor, two anonymous referees, Marcelo Fernandes, Asger Lunde, Carsten Tanggaard, Pedro Valls and seminar participants at several workshops and conferences. We thank for the opportunity to run the computational analysis on the High Performance Computing Cluster supported by the Research and Specialist Computing Support service at the University of East Anglia (UEA).

*Corresponding author at: School of Economics, University of East Anglia, Norwich - UK. E-mail: G.Fruet-Dias@uea.ac.uk

1 Introduction

The proliferation of trading venues within many developed security markets has been a new trend in the financial equity market. The U.S., for instance, features eleven exchanges, eighty-five alternative trading systems (ATS) and more than two-hundred dealers.[1] The market fragmentation phenomenon experienced in the U.S. highlights the importance of the microstructure of trading venues, as traditional listing exchanges have created markets within markets, changed markets' pricing structures, and set up specialized microstructures to attract specific trading clienteles (Menkveld, 2014, O'Hara, 2015, and Menkveld, 2016). Retail brokers have responded to this new market setting by routing their orders to multiple trading platforms (Battalio et al., 2016). Consequently, the way traders and markets learn and disseminate information is also dispersed across markets (O'Hara, 2015).

In this context of a decentralized system in which there are multiple prices for one homogeneous security, the study of information processing among trading venues becomes highly complex and, at the same time, of great significance for both market participants and regulators. In this paper, we take a different angle to the price discovery literature and investigate information processing in the stochastic process driving stock's volatility.[2] The economic reasoning of the volatility discovery framework hinges on the fact that economic agents do not process volatility related news instantaneously and on the relevance of the informational content driving the volatility process. As for the former, recent evidence in Lochstoer and Muir (2022) shows that financial agents have sticky expectations about volatility that generates initial underreaction and delayed overreaction to volatility news. As for the volatility informational content, French and Roll, 1986 document that volatility is mainly caused by private information which affects prices when informed investors trade. Differently from lagged returns, the volatility process contains important information that

[1] https://www.sec.gov/news/statement/us-equity-market-structure.html

[2] The finance and econometrics literature has long viewed volatility as a separate stochastic process (see the excellent survey on stochastic volatility models in Shephard and Andersen, 2009).

produces some degree of predictability in future returns (see, for instance, Ng and Ludvigson 2007 and Bollerslev et al. 2009). More specifically, Bollerslev et al. (2013) find that the implied and realized variation fractionally cointegrate and this long-run equilibrium relationship predicts returns over interdaily horizons. Finally, the information contained in the volatility process has compelled traders to use derivative markets to trade volatility (Ni et al., 2008). All in all, the volatility process reveals information that is not present in lagged returns and hence it is not captured by the price discovery framework. This means that the volatility discovery analysis can be a valuable tool to identify the source (*what* market) of this information.

The price discovery literature has identified the unobserved efficient price by extracting commonalities within the transaction prices of a homogeneous security. In this context, prices are cointegrated, and the vector error correction (VEC) model has become the workhorse of price discovery analyses. We rely on a similar intuition and exploit the evidence that estimates of the integrated variance (realized measures) are long memory processes and cointegrated in order to identify the latent integrated variance of the efficient price using the fractionally cointegrated vector autoregressive (FCVAR) model of Johansen and Nielsen (2012).[3] We then investigate how different markets impound information into the integrated variance of the efficient price.

We make the following methodological and theoretical contributions: First, we build upon the continuous-time theory and characterize the volatility discovery phenomenon, leading to the derivation of the volatility discovery measures. Second, we show that recovering the efficient price by means of a permanent and transitory decomposition (building block of any price discovery measure) generally does not suffice to identify the integrated variance of the efficient price. Third, we resolve the issue of the generated regressor problem

[3] Well-documented evidence is found in the literature to show that estimates of integrated variance (e.g., realized variance) depict long memory features, i.e., these estimates are characterized as highly persistent and presenting an autocorrelation function that decays at a hyperbolic rate (Andersen and Bollerslev, 1997, Andersen et al., 2003, Corsi, 2009, among others).

in the FCVAR framework when using realized measures. We employ double asymptotics to establish the limiting distribution of the volatility discovery measures, enabling correct inferences on the volatility discovery process.

Our empirical application investigates the volatility discovery mechanism for 30 of the most actively traded stocks in the U.S over a large sample period of 7 years. The tick-by-tick quotes consists of firms from different industries that are traded in 3 trading venues. We document that markets indeed trade towards the efficient integrated variance, i.e., market specific integrated variances are fractionally cointegrated. Our results reveal that trading venues incorporate new information into the stochastic volatility process in an individual manner. In line with our theoretical results, the price discovery analysis identifies distinct market leaders than those based on the volatility discovery measures.

Concurrently with this work, Dimpfl and Elshiaty (2021) reformulate our model to discrete time and study volatility discovery in cryptocurrencies markets. More specifically, they employ a stochastic volatility model fitted to daily returns to obtain daily volatility estimates and compute the volatility discovery measures using the FCVAR model. Forte and Lovreta (2019) also apply FCVAR model. They find that the CDS implied equity volatility and a equity options based index (VSTOXX) are fractionally cointegrated. Finally, Baule et al. (2017) deviate from our approach and adapt Hasbrouck's (1995) information share measure to stationary processes to study volatility discovery for options and bank-issued options (warrants).

The remainder of the paper proceeds as follows. Section 2 introduces the volatility discovery framework. Next, Section 3 shows how to identify the integrated variance of the efficient price as a function of both price and volatility discovery measures. Section 4 derives the asymptotic properties of the volatility discovery measures and Section 5 presents an empirical application. Section 6 offers concluding remarks. The Supplementary material presents the proofs, a simulation exercise, and additional empirical results.

2 The Volatility Discovery Framework

Asset price theory postulates that the efficient log-price of a financial security at time t must follow a *semimartingale* process defined in some filtered probability space $(\Omega, \mathcal{F}, \mathcal{P})$ (see, among others, Andersen et al., 2003; Barndorff-Nielsen et al., 2011). Additionally, econometrics and finance literature have widely considered the volatility process of the efficient price a stochastic process, such that the efficient price can be written as

$$m_t = \int_{t_0}^{t} a_u \, \mathrm{d}u + \int_{0}^{t} \sigma_{m,u} \, \mathrm{d}W_u + m_0, \tag{1}$$

where a is a predictable locally bounded drift, σ_m is a càdlàg volatility process denoted as the efficient stochastic volatility process, and, without loss of generality, W is a Brownian motion. It is important to note that m_t is driven by two separate stochastic processes: W and σ_m, which reflect new information arrival to the price and volatility processes, respectively.[4] Furthermore, the efficient price is considered a latent process not only due to trading frictions but also as a result of market fragmentation.

Our framework consists of investigating a homogeneous asset traded in S different markets. In general terms, the $S-$dimensional log-price process is modelled as a *Brownian semimartingale* defined in some filtered probability space $(\Omega, \mathcal{F}, \mathcal{P})$. The $S-$dimensional daily log-price process then reads

$$P_t = \int_{t-1}^{t} \widetilde{a}_u \, \mathrm{d}u + \int_{t-1}^{t} \widetilde{\Sigma}_u \, dW_u + P_0, \tag{2}$$

where $\widetilde{a}$ is a predictable locally bounded drift, $\widetilde{\Sigma}$ is a càdlàg volatility matrix with $\Sigma = \widetilde{\Sigma}\widetilde{\Sigma}'$ denoting the càdlàg spot covariance matrix, and W is a S-dimensional vector of Brownian

[4]The exact parametric functional form of the stochastic volatility process σ_m is not relevant to our analysis, and hence we only stress that σ_m is driven by a different (possibly correlated) Brownian motion from W.

motions. The quadratic variation of P is defined as

$$\langle P_t, P_t' \rangle_{t-1\,:\,t} = \int_{t-1}^{t} \widetilde{\Sigma}_u \widetilde{\Sigma}_u' \, \mathrm{d}u = \int_{t-1}^{t} \Sigma_u \, \mathrm{d}u \tag{3}$$

where $\langle P_t, P_t' \rangle_{t-1\,:\,t} = \operatorname{plim}_{n \longrightarrow \infty} \sum_{i=1}^{n} \left(P_{t_i} - P_{t_{i-1}}\right) \left(P_{t_i} - P_{t_{i-1}}\right)'$ with n denoting the number of intraday observations such that $t - 1 = t_0 < t_1 \ldots < t_n = t$ and T being the total number of trading days, such that $t = 1, \ldots, T$. Because the price process in (2) does not consider jumps, the diagonal elements of $\langle P_t, P_t' \rangle_{t-1\,:\,t}$ are the market-specific integrated variances (IV). The IV is central to the pricing of financial instruments, portfolio allocation, and risk management. In a fragmented market context, not only market prices but also their stochastic volatilities are expected to gather information at different speeds, following market characteristics such as market design, trading costs, trading clientele, liquidity, and the presence of informed traders.

The price discovery literature typically adopts the VEC model to approximate the dynamics of (2), identify the single common stochastic trend seen as the efficient price, and obtain consistent estimates of the price discovery measures. Among the most popular price discovery measures are the Booth et al.'s 1999 component share measure and variants of Hasbrouck's (1995) information share (IS) measure.[5]

To investigate the dynamics of the market-specific IVs, we first acknowledge the fact that these IVs are tied to a long-run equilibrium and track a common latent stochastic process seen as the integrated variance of the efficient price. Second, we need to define feasible estimates of market-specific IVs originated from (3). We generically denote these estimates as realized measures (RM), which are consistent estimators computed with ultra-high-frequency data. There are several consistent estimators for the IV (see Mcaleer and Medeiros, 2008 for a comprehensive survey of realized measures estimators), thus, at this

[5]See Supplementary material for an overview on the price discovery framework.

point, we do not distinguish among these different estimators, apart from requiring them to be consistent estimates of the market-specific IVs. Third, as opposed to security prices, which are known to be integrated of order one, $I(1)$, processes, there is well-documented evidence that RMs are persistent and characterized by a long memory. The long memory feature of RMs follows from early work on the conditional volatility of daily financial returns (Baillie et al., 1996) and on modelling RMs (Andersen and Bollerslev, 1997, Andersen et al., 2003, Corsi, 2009). In particular, Andersen et al. (2003) note that "*The slow hyperbolic autocorrelation decay symptomatic of long memory is evident ...*". This stylized fact of RMs implies that these series are fractionally integrated of some order d. A process z_t is said to be fractionally integrated of order d if $(1-L)^d z_t = \eta_t$, where η_t is an integrated of order zero (stationary) process. It follows that $\Delta^d := (1-L)^d$ is the fractional difference operator defined in terms of the binomial expansion $\Delta^d = \sum_{i=0}^{\infty} \binom{d}{i} (-L)^i = \sum_{i=0}^{\infty} \zeta_{i,d} L^d$.

One way to accommodate these characteristics is to approximate the daily dynamics of RMs using the fractionally cointegrated vector autoregressive (FCVAR) model of Johansen (2008) and Johansen and Nielsen (2012),

$$\Delta^d R_t = \alpha\beta'\Delta^{d-b} L_b R_t + \sum_{j=1}^{\kappa} \Gamma_j \Delta^d L_b^j R_t + \varepsilon_t, \quad t = 1, ..., T, \tag{4}$$

where $R_t = (R_1, \ldots, R_S)' = \log(\mathrm{diag}(\mathbf{R}_t))$ is a $S \times 1$ vector that concatenates the element-wise logarithmic of the daily market-specific RMs, $\mathbf{R}_t$ is a $S \times S$ consistent estimator of $\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u$, $L_b = 1 - (1-L)^b$ is the usual lag operator for fractional processes, d and b are the fractional order of the RMs and the degree of fractional cointegration, respectively, and ε_t is independently and identically distributed (i.i.d.) with a mean zero and positive-definite covariance matrix Φ. The parameters in the FCVAR framework have the same common interpretation as the parameters in the standard VEC model. In that, α corresponds to the speed of adjustment and β is the cointegrating vector so that $\beta' R_t$ is integrated of order

$d-b$ and represents the long-run equilibrium relation. It is relevant to note two important aspects of the approximation in (4). First, because the RMs are not expected to diverge over time, the rank of $\alpha\beta'$ should equal $S-1$ in that the RM's share a single common fractional stochastic trend. We see this stochastic trend as an estimate of the logarithm of the IV of the efficient price. Second, we expect deviations between the RMs in the different markets to be transient, i.e., a short memory (covariance stationary) process driven by the trading-venue-specific characteristics, such as cost structure, different degrees of transparency, the speed of order execution and different trader groups, among other features. This stylized fact is accommodated by assuming $d = b$ in (4), which implies that $\beta' R_t$ is an $I(0)$ process. Finally, we allow the RMs to present a fractional order in the range of $0.5 \leq d < 1$, implying that these measures may be nonstationary but mean reverting.[6] It is possible to extend the FCVAR approximation in (4) to the estimates of the market-specific spot volatilities at a cost of a larger measurement error on their estimates. Nevertheless, our simulation study shows that fitting a FCVAR model to the daily RMs still allows us to capture the partial adjustment dynamics that occurs on the spot volatilities.[7]

Multivariate random walk decompositions are central to constructing any price discovery measure, as they decompose observed prices into two components: the $I(1)$ efficient price – the stochastic trend – and an $I(0)$ process that is associated with the portion of information that has no permanent impact on prices. In the same way that the random walk decomposition constitutes the basic building block of price discovery analysis, the fractional counterpart of the Granger representation theorem serves as a building block for volatility discovery analysis. By decomposing the RMs into a long memory term common to all markets and market-specific $I(0)$ terms, we can identify the efficient price's IV and investigate how different venues incorporate information and adjust to the long-run equilib-

[6]This is consistent with previous empirical findings, such as those of Rossi and de Magistris (2014), among others, who find estimates of d greater than 0.5.

[7]See Supplementary material for the details.

rium. Johansen (2008) and Johansen and Nielsen (2012) provide the fractional counterpart of the Granger representation theorem and decompose the process R_t into $I(d)$ and $I(0)$ components:

$$R_t = \Psi \Delta_+^{-d} \varepsilon_t + X_t, \quad t = 1, ..., T, \tag{5}$$

where $\Psi = \beta_\perp \left[\alpha_\perp' \left(I_S - \sum_{j=1}^{\kappa} \Gamma_j\right) \beta_\perp\right]^{-1} \alpha_\perp'$, Δ_+^{-d} is a truncated version of the fractional difference operator of order d; X_t is an $I(0)$ stochastic term; and $\alpha_\perp$ and $\beta_\perp$ are the orthogonal complements of α and β, respectively, with $\alpha' \alpha_\perp = 0$, $\alpha_\perp' \iota_S = 1$, $\beta' \beta_\perp = 0$, and ι_S denoting a $S \times 1$ vector of ones. Similarly to the price discovery case, if $\beta = (I_{S-1}, -\iota_{S-1})'$ and $\mathrm{rk}(\alpha\beta') = S - 1$, the matrix Ψ has common rows. Section 5.1 provides strong empirical support for the cointegrating vector akin to $\beta = (I_{S-1}, -\iota_{S-1})'$ in the FCVAR setup.[8] Hence, a similar interpretation as that for the price discovery analysis holds in the volatility discovery setup, implying that R_t shares a single fractional stochastic trend that is integrated of order d, given by

$$R_{m,t} = \psi \Delta_+^{-d} \varepsilon_t = \left[\alpha_\perp' \left(I_S - \sum_{j=1}^{\kappa} \Gamma_j\right) \beta_\perp\right]^{-1} \alpha_\perp' \Delta_+^{-d} \varepsilon_t, \quad t = 1, ..., T, \tag{6}$$

where ψ accounts for the common row of Ψ. Notably, $R_{m,t}$ possesses the long-term persistence and slow hyperbolic decay discussed in Andersen et al. (2003), and it can be seen as the natural estimate of the logarithm of the efficient price's IV. Given that $R_{m,t}$ can be recovered by $\psi \Delta_+^{-d} \varepsilon_t$, we can investigate how it is tied to innovations to the RMs of the different markets.

In view of the Granger representation in (6), $\left[\alpha_\perp' \left(I_S - \sum_{j=1}^{\kappa} \Gamma_j\right) \beta_\perp\right]^{-1}$ is essentially a normalization factor, meaning that $\alpha_\perp$ plays the key role on identifying the fractional stochastic trend. In turn, we define $\alpha_\perp$ as a measure of volatility discovery. The elements

[8] Similarly to the orthogonal complements of the parameters in the VEC model, α and β are not unique and hence $\beta_\perp$ and $\alpha_\perp$ are also not fully identified. Without loss of generality, we impose the normalizations $\beta = \iota_S$ and $\alpha_\perp' \iota_S = 1$.

of $\alpha_\perp$ show how innovations from the logarithm of the market-specific RMs contribute to $R_{m,t}$. As the element of $\alpha_\perp$ for a given market increases, the importance of this market for the volatility discovery process increases.

The economic interpretation of $\alpha_\perp$ is therefore given by its ability to identify where (in *what* markets) the informational content driving the efficient volatility is revealed. In the context of highly fragmented markets, $\alpha_\perp$ has the potential to guide what markets should be used when computing the realized measures, meaning that the markets with the highest volatility discovery measures should be preferred to markets with low importance to the volatility discovery mechanism. Moreover, similar to the price discovery analysis, the volatility discovery framework can be a valuable component of a comprehensive toolbox for assessing market quality. By combining the insights from volatility discovery measures with other relevant indicators, we can gain a deeper understanding of market dynamics, information efficiency, and overall market quality. Because the information content of the volatility process has predictive power for future returns and volatility (see, for instance, Patton and Sheppard 2015 and Baltussen et al. 2018), knowing the source of this information has the potential to improve derivatives pricing and out-of-the sample forecasts.

3 Identification of the efficient price's IV

The discussion posed in the previous section relies on discrete-time approximations to introduce the volatility discovery framework. This section assumes a parametric structure to (2) and to the daily dynamics of its quadratic variation, such that we can work out the identification of the IV of the efficient price. Next, we provide a theoretical bridge between the price and volatility discovery measures and disentangle the role played by them on impounding information to the IV of the efficient price. The proofs for the theoretical results are available on the Supplementary material.

Our first goal is to write the IV of the efficient price as a function of both price and volatility discovery measures. To this extent, we need to specify a data-generating process for the general market-specific price process in (2) and the daily dynamics for the market-specific IVs. The price process is such that the drift parameter has reduced rank, i.e., prices cointegrate, and the diffusion matrix is stochastic. Assumption 1 provides the technical conditions.

Assumption 1 The log-prices on any given day of an asset that trades on S multiple venues follow the S-dimensional Brownian semimartingale

$$\mathrm{d}P_t = \gamma\delta' P_t\,\mathrm{d}t + \widetilde{\Sigma}_t\,\mathrm{d}W_t, \tag{7}$$

where γ and δ have dimensions $S \times S-1$, $\mathrm{rk}\,(\gamma\delta') = S-1$, $\delta' = (I_{S-1}, -\iota_{S-1})$, $\delta'\gamma$ is a $S-1 \times S-1$ matrix of rank $S-1$ with all eigenvalues having negative real parts, meaning that the eigenvalues of $\gamma\delta'$ lie in the left open half-plane of the complex plane and at zero, and $\Sigma_t = \widetilde{\Sigma}_t\widetilde{\Sigma}_t'$ is a càdlàg process with $[\Sigma_t]_{s,s'} > 0$ for all $s, s' = 1, \ldots S$, where $[.]_{s,s'}$ denotes the $(s,s')^{\mathrm{th}}$ element of a matrix.

The process in (7) is Markovian and the discrete-time parameters obtained from its discretization coincide with those from the reduced-rank Gaussian Ornstein-Uhlenbeck (OU) process in Kessler and Rahbek (2004). In that, Kessler and Rahbek's (2004) Theorem 1 holds automatically, meaning that the subspace in $\mathbb{R}^{S-1}$ spanned by the columns of the discrete-time speed-of-adjustment parameters is the same of those spanned by γ.

Next, we establish a multivariate random walk decomposition of the process in (7). To decompose (7) into long- and short-run components, we employ the identity $I_S = \delta_\perp\left(\gamma_\perp'\delta_\perp\right)^{-1}\gamma_\perp' + \gamma\left(\delta'\gamma\right)^{-1}\delta'$ where $\delta_\perp = \iota_S$ and $\gamma_\perp'\iota_S = 1$, where $\gamma_\perp$ and $\delta_\perp$ are the orthogonal complements of γ and δ, respectively. The first and second terms in this identity are orthogonal and hence satisfy the Gonzalo and Granger's (1995) definition of permanent and transitory components. More specifically, $(\gamma_\perp' \delta_\perp)^{-1} \gamma_\perp' P_t$ returns the stochastic trend common to all markets, i.e, a common Brownian semi-martingale with initial value given by $(\gamma_\perp' \delta)^{-1} \gamma_\perp P_0$, whereas $\delta' P_t$ is an $I(0)$ component that admits a continuous-time Wold representation. Proposition 1 formalizes the results.

Proposition 1 Let Assumption 1 holds. The permanent and transitory (P-T) decomposition then reads

$$P_t = \delta_\perp f_t + \gamma \left(\delta' \gamma\right)^{-1} Z_t, \tag{8}$$

where $f_t = \gamma_\perp' \left(\int_{t_0}^t \widetilde{\Sigma}_u \, \mathrm{d}W_u + P_0 \right)$ is the stochastic trend and $\delta' P_t = Z_t$.

The P-T decomposition generalizes the results in Kessler and Rahbek (2001) to the case of a stochastic covariance matrix. The common stochastic trend, f_t, driving the price process is a Brownian semi-martingale and $\gamma_\perp$ is used as a price discovery measure (the component share measure). Furthermore, according to Proposition 1, the IV of f_t is given by $IV_{f,t} = \gamma_\perp' \left(\int_{t-1}^t \Sigma_u \, \mathrm{d}u \right) \gamma_\perp$.

Next, we investigate the relationship between the quadratic variation of P_t and $IV_{f,t}$. To this extent, we start by writing the daily quadratic variation of P_t in (7) as [9]

$$\langle P_t, P_t' \rangle_{t-1:t} = \int_{t-1}^t \widetilde{\Sigma}_u \widetilde{\Sigma}_u' \, \mathrm{d}u = \int_{t-1}^t \Sigma_u \, \mathrm{d}u. \tag{9}$$

In the fragmented markets setting and in the absence of jumps in (7), the diagonal elements of (9) are the market-specific IVs, which we denote as $IV_t = (IV_{1,t}, \ldots, IV_{S,t})' = \mathrm{diag}(\langle P_t, P_t' \rangle_{t-1:t})$. The empirical estimators of the quadratic variation usually adopted in the realized measures literature, e.g., the realized variance estimator, are consistent estima-

[9] The drift component is slower moving that the diffusion component, resulting in zero quadratic variation. As a result, the parametrization of the drift in Assumption 1 does not impact the quadratic variation of observed prices, which remains the same as in equation (3).

tors of $\langle P_t, P_t' \rangle_{t-1:t}$. Therefore, in the context of an homogeneous asset traded on multiple markets, the RMs estimate the market-specific IVs (and covariances) rather than the IV of the efficient price. The price discovery literature advocates using f_t from the P-T decomposition to identify the IV of the efficient price. However, this identification strategy has an important caveat: $IV_{f,t}$ only takes into consideration the weights given by $\gamma_\perp$ and hence provides no information about any potential learning and temporal feedback mechanisms embedded into Σ_t in Assumption 1. More concretely, in order to identify the partial adjustments dynamics in the market-specific IVs, it is necessary to specify a parametric data-generating process for these quantities.

As briefly discussed in Section 2, much evidence from both the empirical and theoretical realized measures literature point towards the importance of taking into account the high degree of persistence and the long memory property in the integrated variance and realized measures (see, among others, Andersen et al., 2003, Corsi, 2009, Bollerslev et al. 2016, and Shi and Yu, 2022). When fitting ARFIMA models to the RMs, statistically significant long memory parameter estimates typically fall within the range of $0 < d < 1$. Consequently, RMs can exhibit either stationary or mean-reverting non-stationary processes (Rossi and de Magistris, 2014 and Shi and Yu, 2022). Alternatively, the long memory feature of RMs can be addressed using Corsi's (2009) heterogeneous autoregressive (HAR) model. The HAR family of models are based on a restricted autoregressive (AR) model that captures the persistence of the RMs by considering the RM in the previous day, its average over the previous week, and its average over 22 days. Baillie et al. (2019) reconcile both approaches and estimate an ARFIMA model with the restrictions implied by the HAR model. Their findings demonstrate the benefits of incorporating the long memory parameter into the HAR model to explain higher-order dynamics. To capture the long memory feature of the IV of the efficient price as well as the fact that market-specific IVs do not diverge and are tied to a long-run equilibrium, we assume the dynamics of the daily market-specific logarithmic IVs follow a FCVAR process. Assumption 2 formalizes their data-generating process.

Assumption 2 1. The daily logarithmic IVs of a given asset that trades on multiple venues follow a FCVAR model as

$$\Delta^d \log(IV_t) = \alpha\beta' L_d \log(IV_t) + \varepsilon_t, \quad t = 1, ..., T, \tag{10}$$

where $0 \leq d < 1$, α and β have dimensions $S \times S-1$, $\mathrm{rk}(\alpha\beta') = S-1$, $\alpha\beta' \neq -I_S$, $\det(\alpha'_\perp \beta_\perp) \neq 0$, $IV_t = \mathrm{diag}\left(\langle P_t, P'_t \rangle_{t-1:t}\right) = (IV_{1,t}, \ldots, IV_{S,t})'$, and log(.) denotes the element-wise logarithmic function.[10]

2. The errors ε_t are i.i.d. process with $\mathbb{E}(\varepsilon_t) = 0$, $Var(\varepsilon_t) = \Phi$ and $\mathbb{E}|\varepsilon_t|^8 < \infty$.

3. The matrix of cointegrating vectors is $\beta' = (I_{S-1}, -\iota_{S-1})$, such that $\beta_\perp = \iota_S$.

Assumption 2 ensures that the market-specific IVs cointegrate and share a single common fractional stochastic trend with long memory order d. This common factor corresponds to the logarithm of the IV of the efficient price and reads

$$\log(IV_{m,t}) = \alpha'_\perp \Delta_+^{-d} \varepsilon_t. \tag{11}$$

Alternatively, we could reformulate Assumption (2) in terms of primitive assumptions regarding both the efficient spot volatility and market-specific and efficient price IVs. Specifically, the fractional Ornstein-Uhlenbeck process of Comte and Renault (1998) models the logarithmic of the spot volatility as a long memory process of order d. Consequently, the logarithmic of the spot volatility associated with the efficient price reads

$$\mathrm{d}\log(\sigma_{m,t}) = -\varpi \, \log(\sigma_{m,t})\,\mathrm{d}t + \varphi\,\mathrm{d}W^d_{m,t}, \tag{12}$$

[10]In line with the empirical evidence in Section 5, we simplify the exposition and set $\kappa = 0$ and $d = b$.

where ϖ is the drift parameter, φ is the volatility (volatility of volatility) parameter, and W_m^d is a fractional Brownian motion (fBm).[11,12] It is important to note that the fBm process in (12) does not imply an arbitrage opportunity because the price process in Assumption 1 is a semi-martingale. The logarithmic of the market-specific spot volatility in (12) inherits the long memory property from the fBm process, as its (pseudo) spectral density has a pole in zero that is a function of d: $f_{\sigma_m}(x) \sim cx^{-2d}$ as $x \to 0$, where $c \in \mathbb{R}_+$ and $0 \leq d < 1$. Importantly, Proposition 2.2 in Comte and Renault (1998) show that $\log(\sigma_m)$, W_m^d, and σ_m share the same long memory order d. Additionally, Proposition 1 in Rossi and de Magistris (2014) establishes that the integrated variance also possesses the same long memory degree d as W_m^d, i.e., $IV_{m,t} = \int_{t-1}^{t} \sigma_{m,u}^2 \,\mathrm{d}u$ is a long memory process of order d. This long memory feature of $\int_{t-1}^{t} \sigma_{m,u}^2 \,\mathrm{d}u$ is equivalent to the fractionally integrated property of the IV of the efficient price implied by Assumption 2. Moreover, because the market-specific integrated variances are tied to a long run equilibrium they can then be expressed as a function of $IV_{m,t}$ and an $I(0)$ zero-mean covariance stationary process $\widetilde{X}_t$, such that $IV_{s,t} = IV_{m,t} + \widetilde{X}_{s,t}$ with $s = 1, \dots S$. This specification establishes a connection with two important conditions implied by Assumption 2. Firstly, it satisfies the definition of fractional cointegration with $\beta' = (I_{S-1}, -\iota_{S-1})$, i.e., $IV_{s,t}$ and $IV_{s',t}$, $s, s' = 1, \dots, S$ and $s \neq s'$, are $I(d)$ processes and $IV_{s,t} - IV_{s',t}$ is an $I(0)$ process. Secondly, it is consistent with Johansen and Nielsen's (2012) Granger representation theorem as implied by Assumption 2, as $IV_{s,t}$ and $IV_{s',t}$ share $IV_{m,t}$ as the common fractional stochastic trend. Finally, the main advantage of employing Assumption 2 over the set of primitive assumptions discussed earlier is that it allows for the identification of $IV_{m,t}$ using d and α (and consequently $\alpha_\perp$). Additionally, it does not require any constraints on the parameters to ensure that estimates of $IV_{m,t}$ are positive,

[11] The fBm is defined as $\int_{-\infty}^{0} ((t-u)^d - (-u)^d) \,\mathrm{d}W_u + \int_0^t \frac{(t-u)^d}{\Gamma(1+d)} \,\mathrm{d}W_u$, where Γ is the Gamma function. The fBm process can also be written in terms of the Hurst index H rather than d, such that $d = H - 1/2$.

[12] Comte et al. (2012) proposes an affine class of long memory volatility processes that involve fractional integration of the square root of the spot volatility process, rather than its logarithm as in (12). Additionally, the fractional OU in (12) can be written in terms of $\log(\sigma_{m,t}^2)$ (Rossi and de Magistris, 2014).

as the FCVAR is defined in terms of the logarithm of the market-specific IVs. Next, we characterize the IV of the efficient price as a function of the price and volatility discovery measures, that is $\gamma_\perp$ and $\alpha_\perp$, respectively.

Theorem 1 Let Assumptions 1 and 2 hold. Let $\widetilde{\widetilde{\Xi}} = \gamma\left(\delta'\gamma\right)^{-1}\delta'$, $k(s)$ be a $S \times 1$ vector that has its s^{th} row equal to one and all the remaining entries equal to zero, $\mathbb{K} = \left(\text{vech}\left(\mathbb{K}_1\right), \ldots, \text{vech}\left(\mathbb{K}_S\right)\right)'$ be a $S \times \frac{1}{2}S(S+1)$ matrix with $\mathbb{K}_s$ denoting a $S \times S$ matrix that has the $(s,s)^{\text{th}}$ entry equal to one and all the remaining entries equal to zero, $\mathbb{L}$ be a $\frac{1}{2}S(S+1) \times S^2$ elimination matrix such that, for any symmetric matrix A, $\text{vech}\left(A\right) = \mathbb{L}\text{vec}\left(A\right)$, and $\log(.)$ be the element-wise logarithm. Then,

(i) if $\gamma_\perp \neq \alpha_\perp$,

$$\begin{aligned}\log\left(IV_{m,t}\right) =& k(1)'\mathbb{K}\mathbb{L}\left(\left(\delta_\perp\gamma_\perp'\right) \otimes \left(\delta_\perp\gamma_\perp'\right)\right)\text{vec}\left(\log\left(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u\right)\right)\\ &+ 2\alpha_\perp'\mathbb{K}\mathbb{L}\left(\left(\delta_\perp\gamma_\perp'\right)\otimes\widetilde{\widetilde{\Xi}}\right)\text{vec}\left(\log\left(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u\right)\right)\\ &+ \alpha_\perp'\mathbb{K}\mathbb{L}\left(\widetilde{\widetilde{\Xi}}\otimes\widetilde{\widetilde{\Xi}}\right)\text{vec}\left(\log\left(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u\right)\right) + \text{initial values;}\end{aligned} \tag{13}$$

(ii) if $\gamma_\perp = \alpha_\perp = k(s)$ with $s \in [1, \ldots, S]$,

$$\log\left(IV_{m,t}\right) = k(1)'\mathbb{K}\mathbb{L}\left(\left(\delta_\perp\gamma_\perp'\right) \otimes \left(\delta_\perp\gamma_\perp'\right)\right)\text{vec}\left(\log\left(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u\right)\right) + \text{initial values;} \tag{14}$$

(iii) if $\gamma_\perp = \alpha_\perp$ and $\gamma_\perp, \alpha_\perp \neq k(s)$ with $s \in [1, \ldots, S]$,

$$\begin{aligned}\log\left(IV_{m,t}\right) =& k(1)'\mathbb{K}\mathbb{L}\left(\left(\delta_\perp\gamma_\perp'\right) \otimes \left(\delta_\perp\gamma_\perp'\right)\right)\text{vec}\left(\log\left(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u\right)\right)\\ &+ \alpha_\perp'\mathbb{K}\mathbb{L}\left(\widetilde{\widetilde{\Xi}}\otimes\widetilde{\widetilde{\Xi}}\right)\text{vec}\left(\log\left(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u\right)\right) + \text{initial values.}\end{aligned} \tag{15}$$

Theorem 1.(i) carries two important implications. Firstly, it shows how the price and volatility discovery mechanisms emerge as different sources on impounding information to the IV of the efficient price. Unlike $IV_{f,t}$, $IV_{m,t}$ fully encompasses the informational content and error correction mechanisms presented in Σ_t through $\alpha_\perp$. Consistent with empirical evidence that demonstrate investor trade based on volatility information (e.g., Ni et al. 2008), incorporating $IV_{m,t}$ instead of $IV_{f,t}$ is expected to enhance trading decisions, while leveraging $\alpha_{\perp}$ identifies what are the sources of information impacting $IV_{m,t}$. Secondly, the identification of the IV of the efficient price using f_t critically hinges on the condition $\gamma_{\perp} = \alpha_{\perp} = k(s)$ with $s \in [1, \ldots, S]$ (Theorem 1.(ii)). Failure to satisfy this condition renders any price discovery measure relying on estimates of $IV_{f,t}$ inconsistent. This has significant implications for any variation of Hasbrouck's (1995) IS measure, as such measures give the relative contribution of each market to the IV of the efficient price.

To further appreciate the limitation of using $IV_{f,t}$ instead of $IV_{m,t}$, we shall discuss the case in which the restriction $\gamma_{\perp} = \alpha_{\perp} = k(s)$ or $\gamma_{\perp} = \alpha_{\perp}$ with $s \in [1, \ldots, S]$ with $s \in [1, \ldots, S]$ holds, i.e., Theorem 1.(ii). In the very special case of $\alpha_{\perp} = \gamma_{\perp} = k(s)$ with $s \in [1, \ldots, S]$, a single market is fully responsible for impounding information to both P_t and $IV_{m,t}$, meaning that the remaining exchanges are irrelevant with respect to information processing. Such scenario appears to be rather unlikely to hold in practise, as trading platforms fiercely compete for order flow and high-frequency traders operate across markets (see, for instance, O'Hara, 2015). Second, the two last terms on the right-hand side of (13) would collapse to zero if the diffusion matrix in (7) is constant, which rules out the presence of stochastic volatility altogether. In such case, there is neither intra- nor inter-daily learning with respect to the volatility processes across the different trading platforms.

Another possible scenario consists of examining the case where price and volatility discovery processes do not happen exclusively in one market, but $\gamma_{\perp}'\alpha = 0$ holds. In such case, Theorem 1.(iii) shows that $\log(IV_{m,t})$ remains a function of both price and volatility discovery measures. More specifically, the cross-quadratic variation is zero and $\alpha_{\perp}$ in (15) shows how $\log(IV_{m,t})$ is related to elements of the the daily quadratic variation of the transitory component in Proposition 1.

4 Estimation and Inference

We now turn our attention to the estimation and inference tools for the volatility discovery measures. We extend the Johansen and Nielsen's (2012) results on consistency and the asymptotic normality of the maximum likelihood (MLE) estimator to the context of double asymptotics (mixed case), i.e., where both long span and infill asymptotics hold. This approach is necessary because we need to replace the latent market-specific IVs by their consistent intraday-based estimates in the FCVAR models. Because consistency of these realized measures are established for $n \to \infty$, consistency and asymptotic normality of the FCVAR-based volatility discovery measures require both $n, T \to \infty$. This approach is consistent with previous work on estimating diffusion processes based on discretely observed data in which the asymptotic properties of the estimators depend on both $n, T \to \infty$ (see, for instance, Prakasa Rao 1983 and Tang and Chen 2009). Specifically, in the context of extremum estimators similar to ours, Phillips and Yu (2009) use the feasible central limit theorem for realized variances in Barndorff-Nielsen and Shephard (2002) as the first step in their two-stage estimation procedure to estimate the parameters in the drift function of a diffusion process.

To obtain the log-likelihood function of the FCVAR model in (10), first note that Assumption 2.2 does not imply that the the errors are Gaussian. This means that the estimation strategy based on the Gaussian likelihood in Johansen and Nielsen (2012) can be interpreted in the context of quasi MLE estimation. In addition to Assumption 2, we need Assumption 3 to ensure the limiting behaviour of the likelihood function for the FCVAR model.[13]

Assumption 3 The initial values IV_{-t}, are uniformly bounded with $IV_{-t} \neq 0$ for $0 \leq t \leq T_0$ and $IV_{-t} = 0$ for $t > T_0$.

[13]Assumptions 2 and 3 are equivalent to Assumptions 1-4 in Johansen and Nielsen (2012).

Assumption 3 concerns the initial values in which the likelihood function is conditioned on and essentially determines that $\Delta^d IV_t$ is well-defined for any $0 \leq d \leq 1$. Given Assumptions 2.1, 2.2, and 3, we follow Johansen and Nielsen (2012) and define the Gaussian likelihood function conditional on IV_{-t}, $0 \leq t \leq T_0$ and $T_0 < T$:

$$\ell_T(\widetilde{\lambda}) = -2T^{-1}L_T(\widetilde{\lambda}) = \log\,\det(\widetilde{\Phi}) + \operatorname{tr}(\widetilde{\Phi}^{-1}T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_t\widetilde{\varepsilon}_t'), \tag{16}$$

where $\widetilde{\lambda} = (\widetilde{d}, \operatorname{vec}(\widetilde{\alpha})', \operatorname{vec}(\widetilde{\beta})', \operatorname{vech}(\widetilde{\Phi})')'$ is a $1 + 2S(S-1) + S(S+1)/2 \times 1$ vector of free parameters and $\widetilde{\varepsilon}_t = \Delta^{\widetilde{d}} \log(IV_t) - \widetilde{\alpha}\widetilde{\beta}' L_{\widetilde{d}} \log(IV_t)$ are the corresponding residuals.[14]

The latent nature of the IVs means that we cannot compute the MLE estimator directly from (16). Therefore, we replace $\log(IV_t)$ in (16) by their consistent high-frequency based estimates R_t. Assume prices are observed at regular spaced intervals, such that there is n intraday prices in a given day. Define $\mathbf{R}_t = \sum_{i=1}^{n-1}\left(P_{t_{i+1}} - P_{t_i}\right)\left(P_{t_{i+1}} - P_{t_i}\right)'$ as the realized covariance estimator of $\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u$, such that $R_t = (R_1, \ldots, R_S)' = \log(\operatorname{diag}(\mathbf{R}_t))$ and $\log(.)$ denotes the element-wise logarithm. The feasible conditional log-likelihood function then reads

$$\ell_{n,T}(\widetilde{\lambda}) = -2T^{-1}L_{n,T}(\widetilde{\lambda}) = \log\,\det(\widetilde{\Phi}) + \operatorname{tr}(\widetilde{\Phi}^{-1}T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_{n,t}\widetilde{\varepsilon}_{n,t}'), \tag{17}$$

where $\widetilde{\varepsilon}_{n,t} = \Delta^{\widetilde{d}} R_t - \widetilde{\alpha}\widetilde{\beta}' L_{\widetilde{d}} R_t$. Note that the notation in (17) includes the number of intraday observations n, which serves to highlight the fact that R_t are consistent estimates of the logarithm of the IVs as $n \to \infty$. As a result, $\ell_{n,T}(\widetilde{\lambda})$ also depends on n. More precisely, we impose the following assumptions on the RMs and feasible conditional log-likelihood function.

Assumption 4 1. As $n \to \infty$, $\sqrt{n}\left(\operatorname{vech}(\mathbf{R}_t) - \operatorname{vech}(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u)\right) \xrightarrow{d} \mathcal{N}(0, \Pi_t)$, where Π_t is a positive-definite matrix (Barndorff-Nielsen and Shephard 2004).

[14] Johansen and Nielsen's (2012) Theorem 4 yields that the likelihood function has a strict minimum at the true vector of the parameters and converges uniformly to its deterministic limit.

2. $\sup_{\widetilde{\lambda}\in\Lambda}\left\|\frac{\partial \ell_{n,T}(\widetilde{\lambda})}{\partial \widetilde{\lambda}'}\right\| = O_p(1)$, with $\widetilde{\lambda} \in \Lambda$ and $\|.\|$ denoting the Euclidean norm, and Λ is a compact set.

3. The initial values R_{-t} are uniformly bounded with $R_{-t} \neq 0$ for $0 \leq t \leq T_0$ and $R_{-t} = 0$ for $t > T_0$.

Assumption 4.1 provides information on the rate of convergence of the RMs estimates, which is relevant for establishing the limiting distribution of $\widehat{\alpha}_{\perp}$ in the mixed asymptotics case. The limiting result in Barndorff-Nielsen and Shephard (2004) is obtained under the assumption of "no-leverage", which corresponds to the case where $\widetilde{\Sigma}_t$ is random but independent of the Brownian motion in (7) and achieves the optimal convergence rate. Assumption 4.2 imposes a Lipschitz type smoothness condition to the feasible conditional log-likelihood function. Assumption 4.3 relates to the initial values and it is analogous to Assumption 3.

The estimator we consider is in the class of extremum estimators and hence it is necessary to show that feasible and infeasible conditional log-likelihood functions and their gradients converge uniformly in probability as n, $T \to \infty$. It then follows that, for a fixed $\widetilde{d}$, the conditional MLE estimator based on (17) reduces to a reduced rank regression similar to the standard VEC case. The likelihood function can be simplified by concentrating out the set of parameters $(\text{vec}(\alpha)', \text{vec}(\beta)', \text{vech}(\Phi)'))'$. This allows us to estimate d through the numerical optimization of the profile likelihood function. As a result, $\widehat{\lambda}$ is the MLE estimator of λ. The limiting distribution of the MLE estimator for λ determines the asymptotic distribution underlining the LR-based cointegration rank test, which is used to determine the number of cointegrating vectors in the FCVAR model (see Johansen and Nielsen, 2012 for more details). Finally, we could use the MLE estimates of α to compute the estimates of the volatility discovery measures: $\widehat{\alpha}_{\perp}$. Theorem 2 formalizes the limiting distribution of the volatility discovery measures. To show that, define $\alpha_{\perp}$ as $\alpha_{\perp} = \frac{1}{\varsigma}\left[\iota_S - \alpha\left(\alpha'\alpha\right)^{-1}\alpha'\iota_S\right]$, where $\varsigma = \iota_S' \iota_S - \iota_S' \alpha(\alpha'\alpha)^{-1}\alpha'\iota_S$ such that $\iota_S'\alpha_\perp = 1$ and $\alpha'\alpha_\perp = 0$ hold. Next, note that $\alpha_\perp$ is a vector-valued continuously differentiable function and define $\nabla\alpha_\perp(\alpha) = \frac{\partial \alpha_\perp}{\partial \text{vec}(\alpha)'}$. Also, let $\nabla\ell_{n,T}(\widetilde{\lambda}_1, \widetilde{\lambda}_2) = \frac{\partial \ell_{n,T}(\widetilde{\lambda})}{\partial \widetilde{\lambda}'}\Big|_{\widetilde{\lambda}=\bar{\lambda}}$ and $\nabla\ell_T(\widetilde{\lambda}_1, \widetilde{\lambda}_2) = \frac{\partial \ell_T(\widetilde{\lambda})}{\partial \widetilde{\lambda}'}\Big|_{\widetilde{\lambda}=\bar{\lambda}}$ be the first-order derivative of $\ell_{n,T}(\widetilde{\lambda}_1, \widetilde{\lambda}_2)$ and $\ell_T(\widetilde{\lambda}_1, \widetilde{\lambda}_2)$, respectively, with respect to $\widetilde{\lambda}$ and evaluated at a vector $\bar{\lambda}$ that is between $\widetilde{\lambda}_1$ and $\widetilde{\lambda}_2$, $\mathbb{J} = (0_{S(S-1)\times 1}, I_{S(S-1)}, 0_{S(S-1)\times S(S-1)}, 0_{S(S-1)\times S(S+1)/2})$ be a $S(S-1)\times 1 + 2S(S-1) + S(S+1)/2$ deterministic matrix such that $\text{vec}(\alpha) = \mathbb{J}\lambda$, and $\frac{\partial^2 \ell_T(\widehat{\lambda})}{\partial \widehat{\lambda}\partial \widehat{\lambda}'} \xrightarrow{p} V$.

Theorem 2 Let Assumptions 1, 2.1, 2.2, 3, and 4 hold and suppose $T_0 \geq T^\nu$ with $\nu < 1/2$. If $d > 1/2$, assume also that $\mathbb{E}\,|\varepsilon_t|^{\bar{q}} < \infty$ with $\bar{q} > (d - 1/2)^{-1}$. If $\frac{n}{T} \to \infty$ as $n, T \to \infty$, then

(i.) $\sup_{\widetilde{\lambda}\in\Lambda} \sqrt{T} \,\left|\ell_{n,T}(\widetilde{\lambda}) - \ell_T(\widetilde{\lambda})\right| = O_p\left(\sqrt{\frac{T}{n}}\,\right)$;

(ii.) $\sqrt{T}\,(\widehat{\alpha}_\perp - \alpha_\perp) \xrightarrow{d} \mathcal{N}\left(0, \nabla\alpha_\perp(\alpha)\mathbb{J}V^{-1}\mathbb{J}'\nabla\alpha_\perp(\alpha)'\right).$

The limiting distribution in Theorem 2 allows us to construct single hypothesis tests on the elements of $\alpha_\perp$ and to formulate a Wald test for joint zero constraints on the $\alpha_\perp$. Specifically, the Wald statistic has the usual limiting chi-squared distribution under the null hypothesis.

In the two-market setting, testing single zero restrictions on the elements of $\alpha_\perp$ is equivalent to testing for weak exogeneity of one of the IVs with respect to the cointegrating vector β. This follows because the null hypothesis $\alpha_{\perp,1} = 0$ implies $\alpha_2 = 0$ and the volatility discovery only occurs in market 2. Similarly, the null hypothesis $\alpha_{\perp,2} = 0$ is equivalent to test the null hypothesis $\alpha_1 = 0$ and implies that volatility discovery happens exclusively in market 1. The rejection of both individual null hypotheses ($\alpha_{\perp,1} = 0$ and $\alpha_{\perp,2} = 0$) suggests that volatility discovery occurs in the two markets. Finally, because Assumption 2 implies a FCVAR with zero autoregressive components, testing for weak exogeneity implies testing for Granger causality (strong exogeneity).

In the general S-dimensional market setting, testing multiple joint $S-1$ zero constraints on the elements of $\alpha_\perp$ allows us to examine whether a single market is solely responsible for impounding information into the efficient IV. For instance, if $S = 3$, rejecting both null hypotheses $\alpha_{\perp,1} = \alpha_{\perp,2} = 0$ and $\alpha_{\perp,1} = \alpha_{\perp,3} = 0$, while failing to reject the null hypothesis $\alpha_{\perp,2} = \alpha_{\perp,3} = 0$ imply finding statistical evidence that market 1 is the single driver for the volatility discovery process. However, this procedure neglects the multiple testing issue to correct for Type I error. We address this issue in our empirical application by applying the Holm-Bonferroni procedure. This procedure involves sorting the the multiple tests p-values in ascending order and sequentially comparing them against the significant level scaled by a function of the number of tests.

5 Empirical Application

Our data consist of 30 of the most actively traded assets in the U.S. extracted from the TAQ database. We consider the two major listing exchanges in the U.S. – Nyse (N) and Nasdaq (T) – and one of the most liquid trading venues, Arca (P). This sample selection reflects a balance between capturing relevant market information and avoiding estimation difficulties associated with a large number of markets. All stocks are simultaneously traded in at least two of the three markets covered in this analysis and represent a broad set of industries. We use tick-by-tick quotes from a sample period of 7 years, from January 2007 to December 2013, which captures the implementation of the National Market System regulation (Reg NMS) and the market fragmentation phenomenon.[15] Table S.3 in the Supplementary material displays the details. We compute the daily RMs using Barndorff-Nielsen et al.'s (2011) realized kernel estimator because it provides consistent estimates of the market IV measures under time-dependent and endogenous market microstructure

[15] The Reg NMS in 2007 allows the entry of new trading venues that are linked together and compete for order flow, liquidity, and trades.

noise, irregularly spaced data, and jumps. These features ultimately allow us to take advantage of mid-quotes observed at a tick-by-tick frequency.

5.1 Model Specification

We use the log RMs to estimate bivariate and trivariate FCVAR models.[16] We consider four market combinations: Nasdaq-Arca, Nasdaq-Nyse, Arca-Nyse, and Nasdaq-Arca-Nyse. This choice accommodates assets that are not simultaneously traded on the three exchanges and ultimately yields a richer and broader sample, which is particularly important for handling firms that are listed on Nasdaq but are not traded on Nyse. We choose the lag length, κ, as the minimum value that makes the LM test for serial correlation on the residuals at the first 10 lags non significant at the 5% significance level.[17]

To evaluate whether the RMs at different trading venues are fractionally cointegrated, we implement the sequential likelihood ratio (LR) test for the cofractional rank of Johansen and Nielsen (2012). For each two-market combination, the RMs are fractionally cointegrated and share a single common fractional stochastic trend if $\mathrm{rk}\,(\alpha\beta') = 1$. As for the three-market case, if $\mathrm{rk}\,(\alpha\beta') = 2$, then the RMs share a single fractional stochastic trend. Table 1 presents the p-values associated with the LR cofractional rank test.[18] The null is not rejected at the 5% significance level in all but five stocks, suggesting that the RMs are fractionally cointegrated and share the IV of the efficient price as a common fractional stochastic trend. As for the long memory feature of the RMs, we find that $\widehat{d}$ is statistically significant at 1% significance level and usually greater than 0.5, confirming the RMs' long memory characteristic (nonstationary and mean reverting). In line with

[16]Estimation results for the FCVAR models are obtained using the computer program by Nielsen and Popiel (2014).

[17]In addition to the requirement that residuals should be a white noise process, we choose κ so that the roots of the characteristic polynomials lie outside the transformed unit circle, (see Johansen, 2008 for a theoretical discussion on identification).

[18]We do not report the p-values when the null hypothesis is $\mathrm{rk}\,(\alpha\beta') = 0$ because we strongly reject the null for the 30 assets in all market combinations.

Assumption 2.3, Table 1 provides evidence that $\widehat{\beta} = (1, -1)'$ and $\widehat{\beta} = (I_2, -\iota_2)'$ hold for the two- and three-market combinations respectively, indicating that the market stochastic volatilities are expected to be equal in equilibrium. To appreciate that, recall that β is not uniquely defined, in that we normalize the left-hand $S-1 \times S-1$ block of β' to be the identity matrix and find that the $S-1 \times 1$ right-hand block of β' equals $-\iota_{S-1}$. In summary, Table 1 provides the necessary preliminary results to ensure that Assumption 2 holds and both identification of the IV_m and inference on the volatility discovery measures are valid.

5.2 Volatility Adjustment

We now turn our attention to the parameter estimates that determine the volatility discovery mechanism. We focus on the estimates of volatility discovery measures $\alpha_\perp$, which is normalized such that $\alpha_\perp' \iota_S = 1$.[19] Recall that the market associated with the highest element of $\widehat{\alpha}_\perp$ leads the volatility discovery process. In view of the Granger representation in (6) and Assumption 2, the logarithm of the integrated variance of the efficient price is given by a weighted sum of the innovations in the different market-specific integrated variances, with weights given by the elements of $\widehat{\alpha}_\perp$. In that, it is natural to interpret these weights as non-negative, such that $0 \leq \widehat{\alpha}_{\perp,s} \leq 1$ with $s = 1, \ldots, S$ and $\widehat{\alpha}_\perp' \iota_S = 1$. In finite sample analysis, it is still possible for some elements of the estimates of $\alpha_\perp$ to turn out to be negative. In these cases, we should not reject the null hypothesis that these volatility discovery measures are equal to zero.

Table 2 presents the estimates of $\alpha_\perp$ for each market combination, along with the significance levels (one-sided test) of the null hypothesis that a single market does not contribute to the volatility discovery process, i.e., $\alpha_{\perp,s} = 0$ with $s \in (\mathrm{T}, \mathrm{P}, \mathrm{N})$.[20] For the

[19]Tables S.5 and S.6 in the Supplementary material report the estimates of α and the LM-test for serial correlation in the residuals, respectively.

[20]Because we find $\kappa = 0$ for nearly all stocks (see Table S.4 in the Supplementary material), the hypothesis

three-market setting, we provide the corresponding p-value for the null hypotheses that two volatility discovery measures are equal to zero. These p-values are denoted as $\mathcal{H}_{s,s'}$, where $s, s' \in (\mathrm{T}, \mathrm{P}, \mathrm{N})$ and $s \neq s'$. For example, $\mathcal{H}_{T,P}$ represents the p-value associated with the null hypothesis $\alpha_{\perp,T} = \alpha_{\perp,P} = 0$.

Considering the first market combination, pointwise analyses show that $\widehat{\alpha}_{\perp,P} > \widehat{\alpha}_{\perp,T}$ holds for 19 of the 28 assets. This suggests that Arca is more important than Nasdaq. A more precise analysis is obtained by examining the hypothesis tests derived from Theorem 2. At the 5% significance level, we find that the volatility discovery mechanism occurs exclusively at Arca for 12 of the 28 stocks. This means that we fail to reject the null hypothesis $\alpha_{\perp,T} = 0$, while simultaneously rejecting the null hypothesis $\alpha_{\perp,P} = 0$. Similar analysis offers statistical evidence that the Nasdaq is the single driver for 5 assets only. In order to address the issue of multiple testing, we apply the Holm-Bonferroni correction procedure at the 5% significance level. We find that that Arca and Nasdaq are the single drivers of the volatility discovery process for 9 and 4 stocks, respectively.[21]

Results from the second market combination also suggest that Nasdaq is less important. We find that $\widehat{\alpha}_{\perp,N} > \widehat{\alpha}_{\perp,T}$ holds for 13 of the 19 stocks, indicating that Nyse leads the volatility discovery process. At the 5% significance level, we find statistical evidence that Nyse is the unique contributor for 9 of the 19 assets, which also implies that Nasdaq does not Granger-cause the Nyse for these stocks. Considering the same significance levels, we find that Nasdaq is the sole contributor for 6 stocks only. These findings remain consistent when applying the Holm-Bonferroni correction procedure.

Regarding the Arca-Nyse market combination Arca and Nyse lead in 11 and 10 of the 21 markets, respectively. When applying the hypothesis tests, we find that Arca and Nyse individually drive volatility discovery for 8 and 9 stocks, respectively, at the 5% significance level. Controlling for multiple tests does not alter the qualitative findings.

tests below can be most often interpreted as Granger causality tests.

[21] See Table S.7 in the Supplementary material for the p-values of the volatility discovery measures.

The two-market setting shows Nasdaq as the least important exchange to the volatility discovery process, whereas Nyse appears as the most efficient one. To obtain a more precise answer to which market drives the volatility discovery process, the natural step forward in our analysis is to extend it to a three-market setting. The final set of results in Table 2 displays the results. Pointwise analyses reveal a stronger role of Nyse when compared to the two-market setting. Nyse dominates in 12 out of the 19 stocks, while Arca and Nasdaq leads in only 5 and 2 stocks, respectively. With joint hypothesis tests at the 5% significance level, we reject the null hypotheses $\alpha_{\perp,T} = \alpha_{\perp,N} = 0$ and $\alpha_{\perp,P} = \alpha_{\perp,N} = 0$ and fail to reject the null hypothesis $\alpha_{\perp,T} = \alpha_{\perp,P} = 0$ for 8 of the 19 stocks, indicating that Nyse is the sole contributor to the volatility discovery process. Similarly, our analysis shows that Arca and Nasdaq are the sole contributor for 1 and 0 stocks, respectively. The Holm-Bonferroni correction procedure confirms Nyse as the single driver in 7 stocks.

Overall, our findings in Table 2 indicate that Nasdaq plays a less significant role compared to Nyse, which emerges as the leading exchange in impounding information for the efficient integrated variance. Using the limiting distribution established in Theorem 2, we find statistical evidence suggesting that a single market typically acts as the sole driver of the volatility discovery process.

We further expand our empirical application to investigate the implications of our theoretical framework, which highlights the presence of two distinct sources of information within the price and volatility discovery mechanisms. Having documented the secondary role of Nasdaq for the volatility discovery mechanism, we next examine whether the same feature holds in the price discovery analysis. We follow the price discovery literature and adopt the VEC model as a discrete approximation of the observed prices in the three market combinations discussed previously.[22] Our choice of price discovery measure is the component share (CS), which is invariant to the sampling frequency, meaning that we can

[22] See Supplementary material for an overview of the price discovery analysis and the estimation details.

consistently estimate it from prices sampled at any sampling intervals (Dias et al., 2020). As there is a trade-off between the choice of sampling interval and the size of the standard errors of the CS measures, we use mid-quotes aggregated at 10 seconds to balance out the negative effect of market microstructure noise and the positive effect of having information at the highest possible frequency.

In general, the price discovery results in Table 3 suggest that Nasdaq leads the price discovery mechanism when compared to both Arca and Nyse. We find that Nasdaq leads in 22 of the 28 stocks for the Nasdaq-Arca market combination and 14 of the 19 stocks for the Nasdaq-Nyse market combination. The Arca-Nyse combination shows generally more balanced results, with Arca leading in 13 of the 21 stocks. The results from the Nasdaq-Arca-Nyse market combination are robust to the two-market setting, with Nasdaq leading in 12 of the 19 stocks. In contrast with the volatility discovery analysis, we find that all markets contribute to the price discovery process irrespectively of the market combination, i.e., we reject the null hypotheses $\gamma_{\perp,s} = 0$ for all $s = 1, \ldots, S$ at the 5% significance level for all but two stocks. This result holds when controlling for the multiple testing issue.

Overall, our results suggest that the volatility and price discovery mechanisms do not necessarily occur in the same trading venue. This finding is in line with Theorem 1, which shows that the price and volatility discovery measures capture how markets incorporate news into two distinct efficient stochastic processes: the efficient price and the efficient stochastic volatility.

6 Conclusions

In the current context of market fragmentation, we develop a theoretical framework to investigate how distinct markets contribute to the efficient price's IV. The economic rationale supporting the volatility discovery analysis rests on the premise that the efficient volatility process is a separate stochastic process. Just as the price discovery measure, the volatility discovery analysis contributes to assess market efficiency and investor protection - reliable and transparent information when making investment decisions using volatility.

The volatility discovery framework builds on evidence that the RMs of a homogeneous asset are tied to a long run equilibrium and share a common factor: the efficient stochastic volatility. We exploit RMs' long memory feature and adopt the FCVAR model to identify and estimate how innovations to market volatilities contribute to the efficient volatility process. We show the role played by both price and volatility discovery measures on the identification of the integrated variance of the efficient price. Next, we establish the limiting distribution of the volatility discovery measures and discuss how to make inference on market leadership. Finally, our empirical application corroborate our theoretical findings: RMs are tied to a long run equilibrium; markets incorporate information distinctively; and volatility and price discovery do not necessarily occur in the same trading venue.

SUPPLEMENTARY MATERIAL

Supplementary Material It contains the proofs for Proposition 1 and Theorems 1 and 2. It also presents a review of the price discovery measures, a simulation exercise, data details, and additional empirical results (SuppVD.pdf file). The authors report there are no competing interests to declare.

References

Andersen, T. G. and T. Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The Journal of Finance LII*(3), 975–1005.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 74*(1), 3 – 30.

Baillie, R. T., F. Calonaci, D. Cho, and S. Rho (2019). Long memory, realized volatility and heterogeneous autoregressive models. *Journal of Time Series Analysis 40*(4), 609–628.

Baltussen, G., S. van Bekkum, and B. van der Grient (2018). Unknown unknowns: Uncertainty about risk and stock returns. *Journal of Financial and Quantitative Analysis 53*(4), 1615–1651.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2011). Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics 162*, 149 –169.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(2), 253–280.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Econometric analysis of realised covariation: high frequency based covariance, regression and correlation in financial economics. *Econometrica 72*, 885–925.

Battalio, R., S. A. Corwin, and R. Jennings (2016). Can brokers have it all? on the relation between make-take fees and limit order execution quality. *The Journal of Finance 71*(5), 2193–2238.

Baule, R., B. Frijns, and M. E. Tieves (2017). Volatility discovery and volatility quoting on markets foroptions and warrants. *Journal of Future Markets 28*, 758–774.

Bollerslev, T., D. Osterrieder, N. Sizova, and G. Tauchen (2013). Risk and return: Long-run relations, fractional cointegration, and return predictability. *Journal of Financial Economics 108*(2), 409–424.

Bollerslev, T., A. J. Patton, and R. Quaedvlieg (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics 192*, 1–18.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *Review of Financial Studies 22*(11), 4463–4492.

Booth, G. G., R. W. So, and Y. Tseh (1999). Price discovery in the German equity index derivatives markets. *The Journal of Futures Markets 19*, 619–643.

Comte, F., L. Coutin, and E. Renault (2012). Affine fractional stochastic volatility models. *Annals of Finance 8*, 337–378.

Comte, F. and E. Renault (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance 8*, 291–323.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics 7*, 174–196.

Dias, G. F., M. Fernandes, and C. M. Scherrer (2020, 01). Price discovery in a continuous-time setting. *Journal of Financial Econometrics*.

Dimpfl, T. and D. Elshiaty (2021). Volatility discovery in cryptocurrency markets. *The Journal of Risk Finance 22*, 313–331.

Forte, S. and L. Lovreta (2019). Volatility discovery: Can the cds market beat the equity options market? *Finance Research Letters 28*, 107–111.

French, K. R. and R. Roll (1986). Stock return variances. the arrival of information and the reaction of traders. *Journal of Financial Economics 17*, 5–26.

Gonzalo, J. and C. Granger (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics 13*, 27–35.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The Journal of Finance 50*, 1175–1198.

Johansen, S. (2008). A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory 24*(3), pp. 651–676.

Johansen, S. and M. Ø. Nielsen (2012). Likelihood inference for a fractionally cointegrated vector autoregressive model. *Econometrica 80*(6), 2667–2732.

Kessler, M. and A. Rahbek (2001). Asymptotic likelihood based inference for co-integrated homogenous gaussian diffusions. *Scandinavian Journal of Statistics 28*, 455–470.

Kessler, M. and A. Rahbek (2004). Identification and inference for multivariate cointegrated and ergodic Gaussian diffusions. *Statistical Inference for Stochastic Processes 7*, 137–151.

Lochstoer, L. A. and T. Muir (2022). Volatility expectations and returns. *Journal of Finance 77*, 1055–1096.

Mcaleer, M. and M. C. Medeiros (2008). Realized volatility: A review. *Econometric Reviews 27*, 10–45.

Menkveld, A. J. (2014). High-frequency traders and market structure. *Financial Review 49*(2), 333–344.

Menkveld, A. J. (2016). The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics 8*, 1 – 24.

Ng, S. and S. Ludvigson (2007). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics 83*, 171–222.

Ni, S., J. Pan, and A. Poteshman (2008). Volatility information trading in the option market. *The Journal of Finance 58*, 1059–1091.

Nielsen, M. Ø. and M. K. Popiel (2014). A matlab program and user's guide for the fractionally cointegrated var model. Technical report, QED working paper 1330.

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics 116*(2), 257 – 270.

Patton, A. J. and K. Sheppard (2015, 07). Good Volatility, Bad Volatility: Signed Jumps and The Persistence of Volatility. *The Review of Economics and Statistics 97*(3), 683–697.

Phillips, P. C. and J. Yu (2009). A two-stage realized volatility approach to estimation of diffusion processes with discrete-data. *Journal of Econometrics 150*, 139–150.

Prakasa Rao, B. L. S. (1983). Asymptotic theory for non-linear least squares estimator for diffusion processes. *Statistics: A Journal of Theoretical and Applied Statistics 14*(2), 195–209.

Rossi, E. and P. S. de Magistris (2014). Estimation of long memory in integrated variance. *Econometric Reviews 33*, 785–814.

Shephard, N. and T. G. Andersen (2009). Stochastic volatility: Origins and overview. In *Handbook of Financial Time Series*, pp. 233 – 254. Springer.

Shi, S. and J. Yu (2022). Volatility puzzle: Long memory or antipersistency. *Management Science 0*, 0.

Tang, C. Y. and S. X. Chen (2009). Parameter estimation and bias correction for diffusion processes. *Journal of Econometrics 149*, 65–81.

Table 1: **Cofractional Rank Test and Estimates of the Cointegrating vector**

We report results for 30 assets considering the Nasdaq and Arca, Nasdaq and Nyse, Arca and Nyse, and Nasdaq, Arca and Nyse market combinations. For each two-markets set of results, the first column ($\mathrm{rk}\,(\alpha\beta'))$ displays the number of cointegrating relations used in the estimation of the FCVAR model. Considering the two-market setting, the first column has the p-value of the likelihood-ratio (LR) cofractional rank test when the null hypothesis is $\mathrm{rk}\,(\alpha\beta') \leq 1$. The second column displays the estimates of d and the last column presents the second element of the cointegrating vector, such that $\beta = (1, \beta_2)'$. As for the three-markets set of results, the first column presents the p-value of the likelihood-ratio (LR) cofractional rank test when the null hypothesis is $\mathrm{rk}\,(\alpha\beta') \leq 2$. The second column displays the estimates of d. The third and fourth columns are the free-varying estimates of the normalized cointegrating vector, such that $\beta = (I_2, -\beta_3)'$ with $\beta_3 = (\beta_{3,1}, \beta_{3,2})'$. $*$, $**$, and $***$ denote significance at 1%, 5%, and 10% significance levels."-" implies that the stock is not traded in at least one of the two trading venues during the entire sample.

	Nasdaq-Arca			Nasdaq-Nyse			Arca-Nyse			Nasdaq-Arca-Nyse			
	LR	d	β_2	LR	d	β_2	LR	d	β_2	LR	d	$\beta_{3,1}$	$\beta_{3,2}$
AA	0.38	0.52***	-1.00	0.36	0.52***	-1.00	0.42	0.53***	-1.00	0.35	0.52***	-1.00	-1.00
AAPL	0.99	0.52***	-1.00	-	-	-	-	-	-	-	-	-	-
BAC	0.70	0.60***	-1.00	0.59	0.60***	-1.00	0.71	0.59***	-1.00	0.57	0.59***	-1.00	-1.00
BRKB	0.04	0.52***	-1.01	-	-	-	-	-	-	-	-	-	-
C	-	-	-	-	-	-	0.94	0.59**	-1.00	-	-	-	-
CSCO	0.72	0.54***	-1.00	-	-	-	-	-	-	-	-	-	-
F	0.83	0.56***	-1.00	0.80	0.56***	-1.01	0.80	0.56***	-1.01	0.78	0.56***	-1.01	-1.01
GE	0.51	0.55***	-1.00	0.26	0.54***	-1.00	0.22	0.53***	-1.00	0.15	0.54***	-1.00	-1.00
GM	0.71	0.54***	-1.00	0.73	0.54***	-1.01	0.60	0.54***	-1.01	0.61	0.54***	-1.01	-1.01
GOOG	0.71	0.51***	-1.00	-	-	-	-	-	-	-	-	-	-
HPQ	0.88	0.51***	-1.00	0.89	0.51***	-1.00	0.96	0.51***	-1.00	0.92	0.50***	-1.00	-1.00
IBM	0.99	0.62***	-1.00	0.95	0.60***	-1.00	0.95	0.64***	-1.01	0.46	0.53***	-1.00	-1.00
JCP	0.02	0.49***	-1.01	0.01	0.49***	-1.00	0.01	0.47***	-0.99	0.00	0.48***	-1.00	-0.99
JNJ	0.70	0.53***	-1.00	0.57	0.52***	-1.01	0.46	0.52***	-1.01	0.48	0.52***	-1.01	-1.01
JPM	0.76	0.61***	-1.00	0.74	0.60***	-1.00	0.81	0.59***	-1.00	0.67	0.60***	-1.00	-1.00
KO	0.85	0.67***	-0.99	0.87	0.51***	-1.01	0.85	0.51***	-1.01	0.86	0.50***	-1.01	-1.02
MO	0.16	0.49***	-1.00	0.14	0.49***	-1.02	0.04	0.48***	-1.02	0.07	0.49***	-1.02	-1.02
MRK	0.72	0.51***	-1.00	0.57	0.51***	-1.02	0.56	0.51***	-1.01	0.66	0.50***	-1.02	-1.01
MRVL	0.00	0.44***	-1.00	-	-	-	-	-	-	-	-	-	-
MS	0.86	0.61***	-1.00	-	-	-	-	-	-	-	-	-	-
MSFT	0.92	0.51***	-1.00	-	-	-	-	-	-	-	-	-	-
NOK	0.64	0.62***	-1.00	0.34	0.51***	-1.00	0.69	0.51***	-1.00	0.99	0.71***	-1.00	-1.00
ORCL	0.58	0.51***	-1.00	-	-	-	-	-	-	-	-	-	-
PFE	0.53	0.52***	-1.00	0.48	0.51***	-1.00	0.49	0.51***	-1.00	0.29	0.51***	-1.00	-1.00
PG	0.57	0.51***	-1.00	0.62	0.52***	-1.00	0.47	0.51***	-1.00	0.46	0.51***	-1.00	-1.01
VZ	0.44	0.51***	-1.00	0.39	0.52***	-1.01	0.24	0.51***	-1.00	0.31	0.51***	-1.01	-1.00
WFC	-	-	-	-	-	-	0.58	0.59***	-1.00	-	-	-	-
WMT	0.04	0.48***	-1.00	0.06	0.49***	-1.00	0.01	0.48**	-1.00	0.05	0.49**	-1.00	-1.00
XOM	0.45	0.73***	-1.00	0.44	0.74***	-1.00	0.89	0.71***	-1.00	0.97	0.57***	-1.00	-1.00
YHOO	0.15	0.47***	-1.00	-	-	-	-	-	-	-	-	-	-

Table 2: **Volatility Discovery Measures**

We report the volatility discovery measures for 30 assets considering the Nasdaq and Arca, Nasdaq and Nyse, Arca and Nyse, and Nasdaq, Arca and Nyse market combinations. T, P, and N denote Nasdaq, Arca and Nyse, respectively. The FCVAR parameters are computed using the MLE estimator of Johansen and Nielsen (2012) where $\mathrm{rk}\,(\alpha\beta') = 1$ and $\mathrm{rk}\,(\alpha\beta') = 2$ for the two- and three-market settings, respectively, κ is chosen according to Table S.4, and $d = b$. For each set of results, we report the elements of the orthogonal complement of α, where the subscript identifies the market, e.g, $\alpha_{\perp,T}$ denotes the orthogonal complement associated with the Nasdaq market. As for the Nasdaq-Arca-Nyse set of results, the three last columns report the p-values associated with the null hypothesis that volatility discovery measures of two markets are equal to zero, e.g., $H_{T,P}$ reports the p-value of the null given by $\alpha_{\perp,T} = \alpha_{\perp,P} = 0$. $*$, $**$, and $***$ denote significance at 1%, 5%, and 10% significance levels. "-" implies that the stock is not traded in at least one of the two trading venues during the entire sample.

	Nasdaq-Arca		Nasdaq-Nyse		Arca-Nyse		Nasdaq-Arca-Nyse					
	$\alpha_{\perp,T}$	$\alpha_{\perp,P}$	$\alpha_{\perp,T}$	$\alpha_{\perp,N}$	$\alpha_{\perp,P}$	$\alpha_{\perp,N}$	$\alpha_{\perp,T}$	$\alpha_{\perp,P}$	$\alpha_{\perp,N}$	$\mathcal{H}_{T,P}$	$\mathcal{H}_{T,N}$	$\mathcal{H}_{P,N}$
AA	0.57^{*}	0.43	0.45^{*}	0.55^{**}	0.52^{**}	0.48^{**}	0.25	0.29	0.47^{*}	[0.26]	[0.12]	[0.18]
AAPL	0.85^{*}	0.15	-	-	-	-	-	-	-	-	-	-
BAC	0.74^{*}	0.26	0.88^{***}	0.12	0.73^{**}	0.27	0.72^{*}	0.24	0.04	[0.05]	[0.29]	[0.85]
BRKB	0.34	0.66^{***}					-	-	-	-	-	-
C	-	-	-	-	1.20^{***}	-0.20	-	-	-	-	-	-
CSCO	0.68^{***}	0.32	-	-	-	-	-	-	-	-	-	-
F	-0.10	1.10^{**}	1.08^{***}	-0.08	1.43^{***}	-0.43	0.09	1.34^{**}	-0.43	[0.01]	[0.64]	[0.14]
GE	0.00	1.00^{**}	0.87^{**}	0.13	1.27^{***}	-0.27	0.07	1.29^{**}	-0.36	[0.03]	[0.78]	[0.15]
GM	-0.11	1.11^{***}	0.64^{***}	0.36^{*}	0.91^{***}	0.09	-0.12	1.08^{**}	0.05	[0.02]	[0.96]	[0.04]
GOOG	0.26	0.74^{**}	-	-	-	-	-	-	-	-	-	-
HPQ	-0.09	1.09^{***}	0.06	0.94^{***}	0.36	0.64^{**}	-0.31	0.51	0.80^{**}	[0.56]	[0.09]	[0.00]
IBM	0.43	0.57^{*}	0.45	0.55	0.76^{*}	0.24	0.20	0.28	0.52^{*}	[0.33]	[0.05]	[0.04]
JCP	0.41^{*}	0.59^{**}	0.14	0.86^{***}	0.24	0.76^{**}	0.05	0.21	0.74^{**}	[0.78]	[0.06]	[0.01]
JNJ	0.45^{**}	0.55^{**}	0.29	0.71^{***}	0.24	0.76^{***}	0.20	0.15	0.65^{***}	[0.43]	[0.01]	[0.00]
JPM	0.66^{**}	0.34	0.83^{***}	0.17	0.65^{**}	0.35	0.66^{**}	0.32	0.03	[0.02]	[0.21]	[0.66]
KO	0.18	0.82^{*}	0.26	0.74^{***}	0.20	0.80^{***}	0.10	0.20	0.70^{***}	[0.46]	[0.00]	[0.00]
MO	0.44^{**}	0.56^{**}	0.27	0.73^{***}	0.24	0.76^{***}	0.13	0.20	0.67^{**}	[0.52]	[0.01]	[0.01]
MRK	0.44^{**}	0.56^{**}	0.02	0.98^{***}	0.04	0.96^{***}	-0.04	0.08	0.96^{***}	[0.96]	[0.00]	[0.00]
MRVL	0.22	0.78^{**}	-	-	-	-	-	-	-	-	-	-
MS	0.63^{*}	0.37	-	-	-	-	-	-	-	-	-	-
MSFT	1.15^{***}	-0.15	-	-	-	-	-	-	-	-	-	-
NOK	-0.31	1.31^{*}	0.57^{**}	0.43^{*}	0.55	0.45	-0.39	1.08	0.31	[0.71]	[0.87]	[0.32]
ORCL	0.93^{***}	0.07	-	-	-	-	-	-	-	-	-	-
PFE	0.13	0.87^{**}	0.31	0.69^{***}	0.35	0.65^{**}	-0.06	0.51	0.56^{**}	[0.35]	[0.22]	[0.02]
PG	0.69^{***}	0.31	0.37^{**}	0.63^{***}	0.25	0.75^{***}	0.37	0.00	0.63^{***}	[0.20]	[0.00]	[0.01]
VZ	0.40^{**}	0.60^{***}	0.30^{*}	0.70^{***}	0.24	0.76^{***}	0.17	0.24	0.59^{**}	[0.34]	[0.02]	[0.00]
WFC	-	-	-	-	0.47^{**}	0.53^{**}	-	-	-	-	-	-
WMT	0.29	0.71^{**}	0.49^{**}	0.51^{**}	0.63^{**}	0.37^{*}	0.14	0.57^{*}	0.28	[0.08]	[0.50]	[0.06]
XOM	-1.64	2.64^{***}	0.03	0.97^{*}	1.77^{**}	-0.77	0.00	-0.25	1.25^{***}	[0.80]	[0.00]	[0.00]
YHOO	-0.01	1.01^{***}	-	-	-	-	-	-	-	-	-	-

Table 3: **Price Discovery Measures**

We report price discovery results for 30 assets considering Nasdaq and Arca, Nasdaq and Nyse, Arca and Nyse, and Nasdaq, Arca, and Nyse market combinations. The VEC parameters are computed using the OLS estimator where $\mathrm{rk}\left(\gamma\delta'\right) = S - 1$, $\delta = (I_{S-1}, -\iota_{S-1})'$ with S denoting the number of markets, and the lag length is chosen as the minimum value which makes the LM test for serial correlation on the residuals at lag 15 to be insignificant at the 5% significance level. For the first three sets of results, the first and second column show the estimate of the CS measure associated with the first and second market, respectively. As for the Nasdaq-Arca-Nyse set of results, the first three columns display the price discovery measures for Nasdaq, Arca, and Nyse, respectively. The last three columns report the p-values associated with the null hypothesis that volatility discovery measures of two markets are equal to zero, e.g., $H_{T,P}$ reports the p-value of the null given by $\gamma_{\perp,T} = \gamma_{\perp,P} = 0$. $*$, $**$, and $***$ denote significance at 1%, 5%, and 10% significance levels. "-" implies that the stock is not traded in at least one of the two trading venues during the entire sample.

	Nasdaq-Arca		Nasdaq-Nyse		Arca-Nyse		Nasdaq-Arca-Nyse					
	$\gamma_{\perp,T}$	$\gamma_{\perp,P}$	$\gamma_{\perp,T}$	$\gamma_{\perp,N}$	$\gamma_{\perp,P}$	$\gamma_{\perp,N}$	$\gamma_{\perp,T}$	$\gamma_{\perp,P}$	$\gamma_{\perp,N}$	$\mathcal{H}_{T,P}$	$\mathcal{H}_{T,N}$	$\mathcal{H}_{P,N}$
AA	0.57***	0.43***	0.60***	0.40***	0.54***	0.46***	0.42***	0.27***	0.32***	[0.00]	[0.00]	[0.00]
AAPL	0.51***	0.49***	-	-	-	-	-	-	-	-	-	-
BAC	0.63***	0.37***	0.67***	0.33***	0.62***	0.38***	0.53***	0.21***	0.26***	[0.00]	[0.00]	[0.00]
BRKB	0.71***	0.29***	-	-	-	-	-	-	-	-	-	-
C	-	-	-	-	0.53***	0.47***	-	-	-	-	-	-
CSCO	0.67***	0.33***	-	-	-	-	-	-	-	-	-	-
F	0.55***	0.45***	0.69***	0.31***	0.65***	0.35***	0.41***	0.33***	0.26***	[0.00]	[0.00]	[0.00]
GE	0.70***	0.30***	0.71***	0.29***	0.60***	0.40***	0.58***	0.17***	0.24***	[0.00]	[0.00]	[0.00]
GM	0.50***	0.50***	0.53***	0.47***	0.53***	0.47***	0.32***	0.27***	0.40***	[0.00]	[0.00]	[0.00]
GOOG	0.68***	0.32***	-	-	-	-	-	-	-	-	-	-
HPQ	0.53***	0.47***	0.60***	0.40***	0.59***	0.41***	0.39***	0.33***	0.28***	[0.00]	[0.00]	[0.00]
IBM	0.37**	0.63***	0.37**	0.63**	0.49***	0.51***	0.22***	0.37***	0.41***	[0.00]	[0.00]	[0.00]
JCP	0.45***	0.55***	0.44***	0.56***	0.48***	0.52***	0.27***	0.34***	0.39***	[0.00]	[0.00]	[0.00]
JNJ	0.69***	0.31***	0.58***	0.42***	0.40***	0.60***	0.52***	0.13***	0.35***	[0.00]	[0.00]	[0.00]
JPM	0.59***	0.41***	0.75***	0.25***	0.73***	0.27***	0.52***	0.31***	0.17***	[0.00]	[0.00]	[0.00]
KO	0.52***	0.48***	0.59***	0.41***	0.57***	0.43***	0.37***	0.34***	0.29***	[0.00]	[0.00]	[0.00]
MO	0.50***	0.50***	0.53***	0.47***	0.53***	0.47***	0.34***	0.28***	0.38***	[0.00]	[0.00]	[0.00]
MRK	0.56***	0.44***	0.63***	0.37***	0.59***	0.41***	0.48***	0.32***	0.20***	[0.00]	[0.00]	[0.00]
MRVL	0.60***	0.40***	-	-	-	-	-	-	-	-	-	-
MS	0.51***	0.49***	-	-	-	-	-	-	-	-	-	-
MSFT	0.56***	0.44***	-	-	-	-	-	-	-	-	-	-
NOK	0.42***	0.58***	0.41***	0.59***	0.39***	0.61***	0.27***	0.23***	0.51***	[0.00]	[0.00]	[0.00]
ORCL	0.57***	0.43***	-	-	-	-	-	-	-	-	-	-
PFE	0.51***	0.49***	0.42***	0.58***	0.38***	0.62***	0.27***	0.20***	0.53***	[0.00]	[0.00]	[0.00]
PG	0.22	0.78**	0.47***	0.53**	0.46***	0.54***	0.15***	0.38***	0.47***	[0.00]	[0.00]	[0.00]
VZ	0.70***	0.30***	0.63***	0.37***	0.40***	0.60***	0.54***	0.11***	0.35***	[0.00]	[0.00]	[0.00]
WFC	-	-	-	-	0.44***	0.56***	-	-	-	-	-	-
WMT	0.64***	0.36***	0.64***	0.36***	0.52***	0.48***	0.48***	0.24***	0.28***	[0.00]	[0.00]	[0.00]
XOM	0.45*	0.55**	0.67**	0.33**	0.73***	0.27***	0.42***	0.40***	0.18***	[0.00]	[0.00]	[0.00]
YHOO	0.60***	0.40***	-	-	-	-	-	-	-	-	-	-

Supplementary Material to An Econometric Analysis of Volatility Discovery

Gustavo Fruet Dias*
School of Economics, University of East Anglia

Fotis Papailias
King's Business School, King's College London, London - UK

Cristina Scherrer
Department of Finance – London School of Economics (LSE)

*Corresponding author at: School of Economics, University of East Anglia, Norwich - UK. E-mail: G.Fruet-Dias@uea.ac.uk

S.1 Proofs

Proof of Proposition 1: Multiply (7) by $\gamma_\perp'$ such as $\gamma_\perp' \, \mathrm{d}P_t = \gamma_\perp' \widetilde{\Sigma}_t \, \mathrm{d}W_t$ (see, among others, Stockmarr and Jacobsen, 1994 and Kessler and Rahbek, 2001). Then, there exists a solution for $\mathrm{d}f_t = \gamma_\perp' \widetilde{\Sigma}_t \, \mathrm{d}W_t$ given by $f_t = \gamma_\perp' \left(\int_{t_0}^{t} \widetilde{\Sigma}_u \, \mathrm{d}W_u + P_0 \right)$ which is a local martingale. As for the transitory component, multiply (7) by δ' such that $\mathrm{d}Z_t = \delta' \gamma Z_t \, \mathrm{d}t + \delta' \widetilde{\Sigma}_t \, \mathrm{d}W_t$. Assumption 1 ensures all eigenvalues of $\delta'\gamma$ have negative real parts, meaning that a stationary solution exists, which finalizes the proof.

Proof of Theorem 1: We decompose the daily logarithmic of the quadratic variation of P_t by applying the identity $I_S = \delta_\perp \left(\gamma_\perp' \delta_\perp\right)^{-1} \gamma_\perp' + \gamma \left(\delta'\gamma\right)^{-1} \delta'$ to the element-wise logarithm of (9), such that

$$\begin{aligned}\log\left(\langle P_t, P_t' \rangle_{t-1:t}\right) =& \delta_\perp \gamma_\perp' \left(\log \left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u \right) \right) \gamma_\perp \delta_\perp' + \delta_\perp \gamma_\perp' \left(\log \left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u \right) \right) \widetilde{\widetilde{\Xi}}' \\ &+ \widetilde{\widetilde{\Xi}} \left(\log \left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u \right) \right) \gamma_\perp \delta_\perp' + \widetilde{\widetilde{\Xi}} \left(\log \left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u \right) \right) \widetilde{\widetilde{\Xi}}', \quad \text{(S.1)}\end{aligned}$$

where $\widetilde{\widetilde{\Xi}} = \gamma \left(\delta'\gamma\right)^{-1} \delta'$. The first and last terms in (S.1) are the quadratic variation of the common stochastic trend and $I(0)$ component of (8), respectively, whereas the second and third terms account for the logarithm of the cross-quadratic variation between the $I(1)$ and $I(0)$ terms. The volatility discovery process concerns the information processing on the diagonal elements of $\langle P_t, P_t' \rangle_{t-1:t}$. Let $\mathbb{K} = \left(\mathrm{vech}\left(\mathbb{K}_1\right), \ldots, \mathrm{vech}\left(\mathbb{K}_S\right)\right)'$ be a $S \times \frac{1}{2} S(S+1)$ matrix with $\mathbb{K}_s$ denoting a $S \times S$ matrix that has the $(s,s)^{\text{th}}$ entry equal to one and all the remaining entries equal to zero, and $\mathbb{L}$ be a $\frac{1}{2}S(S+1) \times S^2$ elimination matrix such that, for any symmetric matrix A, $\text{vech}(A) = \mathbb{L}\text{vec}(A)$. It then follows that

$$\begin{aligned}\log(IV_t) \quad &= \mathbb{KL}\left((\delta_\perp \gamma_\perp') \otimes (\delta_\perp \gamma_\perp')\right) \text{vec}\left(\log\left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u\right)\right) \\ &\quad + 2\mathbb{KL}\left((\delta_\perp \gamma_\perp') \otimes \widetilde{\Xi}\right) \times \text{vec}\left(\log\left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u\right)\right) \\ &\quad + \mathbb{KL}\left(\widetilde{\Xi} \otimes \widetilde{\Xi}\right) \text{vec}\left(\log\left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u\right)\right), \end{aligned} \tag{S.2}$$

where $\log(IV_t) = \mathbb{KL}\text{vec}\left(\log\left(\langle P_t, P_t' \rangle_{t-1:t}\right)\right)$ is a $S \times 1$ vector collecting the diagonal elements of $\log\left(\langle P_t, P_t' \rangle_{t-1:t}\right)$, i.e., the market-specific IVs.[1] Next, in order to relate $\log(IV_t)$ with the logarithm of the IV of the efficient price, multiply both sides of (10) by $\alpha_\perp'$ to obtain $\alpha_\perp' \Delta^d \log(IV_t) = \alpha_\perp' \varepsilon_t$. Further multiplication of both sides by Δ_+^{-d}, using $\Delta_+^{-d}\Delta^d = 1_+ + \Delta_+^{-d}\Delta_-^d$, and employing the common fractionally stochastic trend in (11), $IV_{m,t} = \alpha_\perp' \Delta_+^{-d} \varepsilon_t$, allows us to rearrange the terms as follows:

$$\log(IV_{m,t}) = \alpha_\perp' \log(IV_t) + \left\{\alpha_\perp' \Delta_+^{-d} \Delta_-^d \log(IV_t)\right\}, \tag{S.3}$$

where the last term on the right-hand side of (S.3) is a linear combination of infinitely many initial values. Finally, substitute (S.2) into (S.3), such that the IV of the efficient

[1] The equality $\left((\delta_\perp \gamma_\perp') \otimes \widetilde{\Xi}\right) = \left(\widetilde{\Xi} \otimes (\delta_\perp \gamma_\perp')\right)$ is not a general rule for the Kronecker product. However, this equality holds in (S.2) due to the specific restrictions in $\delta_\perp \gamma_\perp'$ and $\widetilde{\Xi}$ implied by γ and δ.

price reads:

$$
\begin{aligned}
\log\left(IV_{m,t}\right) &= \alpha_\perp' \mathbb{KL}\left(\left(\delta_\perp \gamma_\perp'\right) \otimes \left(\delta_\perp \gamma_\perp'\right)\right) \operatorname{vec}\left(\log\left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u\right)\right) \\
&\quad + 2\alpha_\perp' \mathbb{KL}\left(\left(\delta_\perp \gamma_\perp'\right) \otimes \widetilde{\widetilde{\Xi}}\right) \operatorname{vec}\left(\log\left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u\right)\right) \\
&\quad + \alpha_\perp' \mathbb{KL}\left(\widetilde{\widetilde{\Xi}} \otimes \widetilde{\widetilde{\Xi}}\right) \operatorname{vec}\left(\log\left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u\right)\right) + \text{initial values}, \\
\log\left(IV_{m,t}\right) &= k(1)' \mathbb{KL}\left(\left(\delta_\perp \gamma_\perp'\right) \otimes \left(\delta_\perp \gamma_\perp'\right)\right) \operatorname{vec}\left(\log\left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u\right)\right) \\
&\quad + 2\alpha_\perp' \mathbb{KL}\left(\left(\delta_\perp \gamma_\perp'\right) \otimes \widetilde{\widetilde{\Xi}}\right) \operatorname{vec}\left(\log\left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u\right)\right) \\
&\quad + \alpha_\perp' \mathbb{KL}\left(\widetilde{\widetilde{\Xi}} \otimes \widetilde{\widetilde{\Xi}}\right) \operatorname{vec}\left(\log\left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u\right)\right) + \text{initial values}, \qquad \text{(S.4)}
\end{aligned}
$$

which concludes the proof of item (i) in Theorem 1.

As for Theorem 1-(ii), first note that if $\gamma_\perp = k(s)$ with $s \in [1, \ldots, S]$, then the s^{th} row of γ is $0_{1\times S-1}$. In that, without loss a generalization, let $\gamma_\perp = k(1)$. We start showing that the second term in the right-hand side of (S.4) is equal to zero. To this extent, write

$$
\begin{aligned}
\alpha_\perp' \mathbb{KL}\left(\left(\delta_\perp \gamma_\perp'\right) \otimes \widetilde{\widetilde{\Xi}}\right) &= \alpha_\perp' \mathbb{KL}\left(\left(\iota_S, \ 0_{S\times S-1}\right) \otimes \widetilde{\widetilde{\Xi}}\right) \\
\alpha_\perp' \mathbb{KL}\left(\left(\delta_\perp \gamma_\perp'\right) \otimes \widetilde{\widetilde{\Xi}}\right) &= \alpha_\perp' \mathbb{KL} \begin{pmatrix} \widetilde{\widetilde{\Xi}} & 0_{S\times S(S-1)} \\ \widetilde{\widetilde{\Xi}} & 0_{S\times S(S-1)} \\ \vdots & \vdots \\ \widetilde{\widetilde{\Xi}} & 0_{S\times S(S-1)} \end{pmatrix} = \alpha_\perp' \left(\widetilde{\widetilde{\Xi}} \ \ 0_{S\times S(S-1)}\right). \qquad \text{(S.5)}
\end{aligned}
$$

Because $\alpha_\perp = \gamma_\perp$, it follows that $\alpha_\perp' \widetilde{\widetilde{\Xi}} = \alpha_\perp' \gamma \left(\delta' \gamma\right)^{-1} \delta' = 0$, which implies that $2\alpha_\perp' \mathbb{KL}\left(\left(\delta_\perp \gamma_\perp'\right) \otimes \widetilde{\widetilde{\Xi}}\right) \operatorname{vec}\left(\log\left(\int_{t-1}^{t} \Sigma_u \, \mathrm{d}u\right)\right) = 0$. Next, we show that $\alpha_\perp' \mathbb{KL}\left(\widetilde{\widetilde{\Xi}} \otimes \widetilde{\widetilde{\Xi}}\right)$ in (S.4) is zero. To see this, note that if $\gamma_\perp = k(1)$, we can partition γ such that $\gamma = \left(0_{S-1\times 1}, \ \bar{\gamma}\right)'$ with $\bar{\gamma}$ denoting the lower $S-1 \times S-1$ transpose part of γ. Using the properties of partitioned matrices, we rewrite $\left(\widetilde{\Xi}\otimes\widetilde{\Xi}\right)$ as

$$\left(\widetilde{\Xi}\otimes\widetilde{\Xi}\right)=\left(\begin{pmatrix}0_{1\times S}\\ \bar{\gamma}\left(\delta'\gamma\right)^{-1}\delta'\end{pmatrix}\otimes\widetilde{\Xi}\right)=\begin{pmatrix}0_{S\times S^2}\\ \left(\bar{\gamma}\left(\delta'\gamma\right)^{-1}\delta'\right)\otimes\widetilde{\Xi}\end{pmatrix}. \tag{S.6}$$

Next, recall that $\mathbb{KL}$ is a $S\times S^2$ selection matrix which selects the diagonal elements of $\widetilde{\Xi}\left(\log\left(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u\right)\right)\widetilde{\Xi}'$ from $\operatorname{vec}\left(\widetilde{\Xi}\left(\log\left(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u\right)\right)\widetilde{\Xi}'\right)$. More specifically, it reads

$$\mathbb{KL}=\begin{pmatrix}k(1)' & 0_{1\times S} & 0_{1\times S} & \dots & 0_{1\times S}\\ 0_{1\times S} & k(2)' & 0_{1\times S} & \dots & 0_{1\times S}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 0_{1\times S} & 0_{1\times S} & 0_{1\times S} & \dots & k(S)'\end{pmatrix}=\begin{pmatrix}k(1)' & 0_{1\times S(S-1)}\\ 0_{(S-1)\times S} & \overline{\mathbb{KL}}\end{pmatrix}, \tag{S.7}$$

where $\overline{\mathbb{KL}}$ denotes the lower right-hand $S-1\times S(S-1)$ block of $\mathbb{KL}$. Using (S.6), (S.7), and $\alpha_\perp=k(1)$, it follows that $\alpha_\perp'\mathbb{KL}\left(\widetilde{\Xi}\otimes\widetilde{\Xi}\right)$ reads

$$\alpha_\perp'\mathbb{KL}\left(\widetilde{\Xi}\otimes\widetilde{\Xi}\right)=\alpha_\perp'\begin{pmatrix}0_{S\times S^2}\\ \overline{\mathbb{KL}}\left(\left(\bar{\gamma}\left(\delta'\gamma\right)^{-1}\delta'\right)\otimes\widetilde{\Xi}\right)\end{pmatrix}=0, \tag{S.8}$$

which finalizes the proof of Theorem 1-(ii).

As for Theorem 1-(iii), start by writing

$$\alpha_\perp'\mathbb{KL}\left(\left(\delta_\perp\gamma_\perp'\right)\otimes\widetilde{\Xi}\right)=\alpha_\perp'\mathbb{KL}\begin{pmatrix}\gamma_\perp'\otimes\widetilde{\Xi}\\ \gamma_\perp'\otimes\widetilde{\Xi}\\ \vdots\\ \gamma_\perp'\otimes\widetilde{\Xi}\end{pmatrix}=\alpha_\perp'\left(\gamma_{\perp,1}\widetilde{\Xi},\ \gamma_{\perp,2}\widetilde{\Xi},\ \dots\ ,\gamma_{\perp,S}\widetilde{\Xi}\right). \tag{S.9}$$

If $\gamma_\perp=\alpha_\perp$ and $\gamma_\perp,\alpha_\perp\neq k(s)$ with $s\in[1,\dots,S]$, $\alpha_\perp'\gamma=\alpha_\perp'\widetilde{\Xi}=0$ holds. It then follows that $\alpha'_\perp \mathbb{K}\mathbb{L}\left((\delta_\perp \gamma'_\perp)\otimes \widetilde{\Xi}\right)=0$. Next, using the properties of the trace operator in the third term on the right-hand side of (S.4), it follows that

$$\alpha'_\perp \mathbb{K}\mathbb{L}\left(\widetilde{\Xi}\otimes\widetilde{\Xi}\right)\mathrm{vec}\left(\log\left(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u\right)\right)=\mathrm{vec}\,(\mathbb{L}'\mathbb{K}')'\left(\alpha_\perp\otimes\widetilde{\Xi}\otimes\widetilde{\Xi}\right)\mathrm{vec}\left(\log\left(\int_{t-1}^{t}\Sigma_u\,\mathrm{d}u\right)\right).$$

It is evident that $\left(\alpha_\perp\otimes\widetilde{\Xi}\otimes\widetilde{\Xi}\right)$ is not zero, which concludes the proof of Theorem 1.

Proof of Theorem 2: (i.) First, using the properties of the trace operator and the positive definiteness of $\widetilde{\Phi}$, $T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_t\widetilde{\varepsilon}'_t$, and $T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_{n,t}\widetilde{\varepsilon}'_{n,t}$, we have

$$\begin{aligned}
\sup_{\widetilde{\lambda}\in\Lambda}\sqrt{T}\,\left|\ell_{n,T}\left(\widetilde{\lambda}\right)-\ell_T\left(\widetilde{\lambda}\right)\right| &= \sup_{\widetilde{\lambda}\in\Lambda}\left\{\sqrt{T}\,\left|\mathrm{tr}\left(\widetilde{\Phi}^{-1}T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_t\widetilde{\varepsilon}'_t\right)-\mathrm{tr}\left(\widetilde{\Phi}^{-1}T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_{n,t}\widetilde{\varepsilon}'_{n,t}\right)\right|\right\}\\
&\leq \sup_{\widetilde{\lambda}\in\Lambda}\left\{\sqrt{T}\,\left|\mathrm{tr}\left(\widetilde{\Phi}^{-1}\otimes\left(T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_t\widetilde{\varepsilon}'_t-T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_{n,t}\widetilde{\varepsilon}'_{n,t}\right)\right)\right|\right\}\\
&\leq \sup_{\widetilde{\lambda}\in\Lambda}\left\{\sqrt{T}\,\left|\mathrm{tr}\left(\widetilde{\Phi}^{-1}\right)\mathrm{tr}\left(T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_t\widetilde{\varepsilon}'_t-T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_{n,t}\widetilde{\varepsilon}'_{n,t}\right)\right|\right\}\\
&\leq \sup_{\widetilde{\lambda}\in\Lambda}\left\{\left|\mathrm{tr}\left(\widetilde{\Phi}^{-1}\right)\right|\right\}\ \sup_{\widetilde{\lambda}\in\Lambda}\left\{\sqrt{T}\,\left|\mathrm{tr}\left(T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_t\widetilde{\varepsilon}'_t\right.\right.\right.\\
&\qquad\left.\left.\left.-T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_{n,t}\widetilde{\varepsilon}'_{n,t}\right)\right|\right\}. \qquad \text{(S.10)}
\end{aligned}$$

Since $\sup_{\widetilde{\lambda}\in\Lambda}\{|\mathrm{tr}(\widetilde{\Phi}^{-1})|\}$ is $O(1)$ according to Assumption 2, we can focus on bounding the second term on the right-hand side of (S.10)

$$\begin{aligned}
\sup_{\widetilde{\lambda}\in\Lambda}\left\{\sqrt{T}\,\left|\mathrm{tr}\left(T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_t\widetilde{\varepsilon}'_t-T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_{n,t}\widetilde{\varepsilon}'_{n,t}\right)\right|\right\} &\leq \sup_{\widetilde{\lambda}\in\Lambda}\left\{\sqrt{T}\,\left|\mathrm{tr}\left(T^{-1}\sum_{t=1}^{T}\left(\widetilde{\varepsilon}_{n,t}-\widetilde{\varepsilon}_t\right)\widetilde{\varepsilon}'_{n,t}\right.\right.\right.\\
&\qquad\left.\left.\left.+T^{-1}\sum_{t=1}^{T}\widetilde{\varepsilon}_{,t}\left(\widetilde{\varepsilon}'_t-\widetilde{\varepsilon}'_{n,t}\right)\right)\right|\right\}. \qquad \text{(S.11)}
\end{aligned}$$

We establish that the right-hand side of (S.11) is $o_p(1)$ in two steps. In the first step, we show that $|\ell_{n,T}(\widetilde{\lambda})-\ell_T(\widetilde{\lambda})|$ converges in probability to zero point-wise. Assumption 4 pertains to the infill dimension and states that $\text{diag}(\mathbf{R}_t) - IV_t = O_p(n^{-1/2})$ as $n \to \infty$ for all $t = 1, \ldots, T$, which also implies that $\widetilde{\varepsilon}_{n,t} - \widetilde{\varepsilon}_t = O_p(n^{-1/2})$ for all $t = 1, \ldots, T$. Next, we consider the time span of the data, denoted by T, in our analysis. We use this to derive the following result: $\sqrt{T}\ |\text{tr}(T^{-1}\sum_{t=1}^{T}(\widetilde{\varepsilon}_{n,t} - \widetilde{\varepsilon}_t)\widetilde{\varepsilon}'_{n,t})| = O_p\left(\sqrt{\frac{T}{n}}\right)$, which implies that $\sqrt{T}\ |\ell_{n,T}(\widetilde{\lambda}) - \ell_T(\widetilde{\lambda})| \xrightarrow{p} 0$ as $\frac{n}{T} \to \infty$ with n, $T \to \infty$. In the second step, we establish uniform convergence, i.e., $\sup_{\widetilde{\lambda}\in\Lambda} \sqrt{T}\ |\ell_{n,T}(\widetilde{\lambda}) - \ell_T(\widetilde{\lambda})| \xrightarrow{p} 0$. Theorem 21.10 in Davidson, 1994, p.337 establishes that Assumption 4.2 is sufficient to ensure that $\ell_{n,T}(\widetilde{\lambda})$ is stochastically equicontinuous for all $\widetilde{\lambda} \in \Lambda$. It then follows by Lemma 2.8 in Newey and McFadden (1994) that $\sup_{\widetilde{\lambda}\in\Lambda} \sqrt{T}\ |\ell_{n,T}(\widetilde{\lambda}) - \ell_T(\widetilde{\lambda})| = O_p\left(\sqrt{\frac{T}{n}}\right)$ as n, $T \to \infty$ and $\frac{n}{T} \to \infty$. This result, in conjunction with Assumptions 1, 2.1, 2.2, 3, and 4, implies that Johansen and Nielsen's (2012) Theorems 4 and 5 hold. Specifically, it establishes that the log-likelihood function based on the realized measures $\ell_{n,T}(\widetilde{\lambda})$ has a strict minimum at λ, and that the maximum likelihood estimator (MLE) obtained from it is consistent.

(ii.) First, let $\|A\|$ denote the Euclidean norm of a matrix, vector, or scalar A, and consider any arbitrary pair of points $\widetilde{\lambda}_1, \widetilde{\lambda}_2 \in \Lambda$. Using the mean value theorem, rewrite $\ell_{n,T}(\widetilde{\lambda}_1) - \ell_{n,T}(\widetilde{\lambda}_2) = \nabla\ell_{n,T}(\widetilde{\lambda}_1, \widetilde{\lambda}_2)(\widetilde{\lambda}_1 - \widetilde{\lambda}_2)$ and $\ell_T(\widetilde{\lambda}_1) - \ell_T(\widetilde{\lambda}_2) = \nabla\ell_T(\widetilde{\lambda}_1, \widetilde{\lambda}_2)(\widetilde{\lambda}_1 - \widetilde{\lambda}_2)$, where $\nabla\ell_{n,T}(\widetilde{\lambda}_1, \widetilde{\lambda}_2) = \frac{\partial \ell_{n,T}(\widetilde{\lambda})}{\partial \widetilde{\lambda}'}\Big|_{\widetilde{\lambda}=\bar{\lambda}}$, $\nabla\ell_T(\widetilde{\lambda}_1, \widetilde{\lambda}_2) = \frac{\partial \ell_T(\widetilde{\lambda})}{\partial \widetilde{\lambda}'}\Big|_{\widetilde{\lambda}=\bar{\lambda}}$, and $\bar{\lambda}$ is vector between $\widetilde{\lambda}_1$ and $\widetilde{\lambda}_2$. It then follows that

$$\begin{aligned}\left[\nabla\ell_{n,T}\left(\widetilde{\lambda}_1, \widetilde{\lambda}_1\right) - \nabla\ell_T\left(\widetilde{\lambda}_1, \widetilde{\lambda}_2\right)\right]\left(\widetilde{\lambda}_1 - \widetilde{\lambda}_2\right) &= \left[\ell_{n,T}\left(\widetilde{\lambda}_1\right) - \ell_T\left(\widetilde{\lambda}_1\right)\right] \\ &\quad - \left[\ell_{n,T}\left(\widetilde{\lambda}_2\right) - \ell_T\left(\widetilde{\lambda}_2\right)\right]. \end{aligned} \tag{S.12}$$

It follows from (S.12) that there exist a vector of constants C_1 such that $C_2 = \left(\left(\widetilde{\lambda}_1 - \widetilde{\lambda}_2\right) C_1'\right)^{-1}$

is well defined. Next, use C_1 and C_2 to bound $\nabla\ell_{n,T}(\widetilde{\lambda}_1,\widetilde{\lambda}_1)-\nabla\ell_T(\widetilde{\lambda}_1,\widetilde{\lambda}_2)$ as

$$\begin{aligned}
\sup_{\widetilde{\lambda}_1,\widetilde{\lambda}_2\in\Lambda}\sqrt{T}\ \left\{\left\|\nabla\ell_{n,T}\left(\widetilde{\lambda}_1,\widetilde{\lambda}_1\right)-\nabla\ell_T\left(\widetilde{\lambda}_1,\widetilde{\lambda}_2\right)\right\|\right\} &\leq \sup_{\widetilde{\lambda}_1,\widetilde{\lambda}_2\in\Lambda}\sqrt{T}\ \left\{\left\|\left(\ell_{n,T}\left(\widetilde{\lambda}_1\right)-\ell_T\left(\widetilde{\lambda}_1\right)\right)\right.\right.\\
&\quad\left.\left.-\left(\ell_{n,T}\left(\widetilde{\lambda}_2\right)-\ell_T\left(\widetilde{\lambda}_2\right)\right)\right\|\left\|C_1'C_2^{-1}\right\|\right\},\\
\sup_{\widetilde{\lambda}_1,\widetilde{\lambda}_2\in\Lambda}\sqrt{T}\ \left\{\left\|\nabla\ell_{n,T}\left(\widetilde{\lambda}_1,\widetilde{\lambda}_1\right)-\nabla\ell_T\left(\widetilde{\lambda}_1,\widetilde{\lambda}_2\right)\right\|\right\} &\leq \sup_{\widetilde{\lambda}_1,\widetilde{\lambda}_2\in\Lambda}\sqrt{T}\ \left\{\left\|\ell_{n,T}\left(\widetilde{\lambda}_1\right)-\ell_T\left(\widetilde{\lambda}_1\right)\right\|\right.\\
&\quad\left.\times\left\|C_1'C_2^{-1}\right\|\right\}+\sup_{\widetilde{\lambda}_1,\widetilde{\lambda}_2\in\Lambda}\sqrt{T}\ \left\{\left\|\ell_{n,T}\left(\widetilde{\lambda}_2\right)\right.\right.\\
&\quad\left.\left.-\ell_T\left(\widetilde{\lambda}_2\right)\right\|\left\|C_1'C_2^{-1}\right\|\right\}. \qquad (S.13)
\end{aligned}$$

Because $\left\|C_1'C_2^{-1}\right\|=O(1)$, it then follows from (i.) that $\sup_{\widetilde{\lambda}_1,\widetilde{\lambda}_2\in\Lambda}\sqrt{T}\ \{|\ell_{n,T}(\widetilde{\lambda}_1)-\ell_T(\widetilde{\lambda}_1)|$ $|C_1'C_2^{-1}|\}$ and $\sup_{\widetilde{\lambda}_1,\widetilde{\lambda}_2\in\Lambda}\sqrt{T}\ \{|\ell_{n,T}(\widetilde{\lambda}_2)-\ell_T(\widetilde{\lambda}_2)||C_1'C_2^{-1}|\}$ have order $O_p\left(\sqrt{\frac{T}{n}}\right)$ in probability. These imply that

$$\sup_{\widetilde{\lambda}_1,\widetilde{\lambda}_1\in\Lambda}\sqrt{T}\ \{|\nabla\ell_{n,T}(\widetilde{\lambda}_1,\widetilde{\lambda}_1)-\nabla\ell_T(\widetilde{\lambda}_1,\widetilde{\lambda}_2)|\}=O_p\left(\sqrt{\frac{T}{n}}\right) \qquad (S.14)$$

Next, rewrite the score function $\nabla\ell_{n,T}(\widehat{\lambda})=\left[\nabla\ell_{n,T}(\widehat{\lambda})-\nabla\ell_T(\widehat{\lambda})\right]+\nabla\ell_T(\widehat{\lambda})$. From (S.13), $\left[\nabla\ell_{n,T}(\widehat{\lambda})-\nabla\ell_T(\widehat{\lambda})\right]$ has order $O_p\left(\sqrt{\frac{T}{n}}\right)$, which implies that $\nabla\ell_{n,T}(\widehat{\lambda})=\nabla\ell_T(\widehat{\lambda})+O_p\left(\sqrt{\frac{T}{n}}\right)$ as n, $T\to\infty$ and $\frac{n}{T}\to\infty$. It then follows that the limiting distribution of the score function given in Johansen and Nielsen's (2012) Lemma 8 remains valid for $\nabla\ell_{n,T}(\widehat{\lambda})$. Next, we apply the results in Johansen and Nielsen's (2012) Theorem 10, such that, as n, $T\to\infty$ and $\frac{n}{T}\to\infty$,

$$\sqrt{T}\left(\mathrm{vec}\left(\widehat{\alpha}\right)-\mathrm{vec}\left(\alpha\right)\right)\ \xrightarrow{d}\mathcal{N}\left(0,\mathbb{J}V^{-1}\mathbb{J}'\right),$$

where $\frac{\partial^2\ell_T(\widehat{\lambda})}{\partial\widehat{\lambda}\partial\widehat{\lambda}'}\xrightarrow{p}V$ and $\mathbb{J}=(0_{S(S-1)\times 1},I_{S(S-1)},0_{S(S-1)\times S(S-1)},0_{S(S-1)\times S(S+1)/2})$ is a $S(S-1)\times 1+2S(S-1)+S(S+1)/2$ deterministic matrix that selects the elements of $\mathrm{vec}(\widehat{\alpha})$

and vec(α) from $\widehat{\lambda}$ and λ, respectively. Note that if $d = b$ and $d > 1/2$, then $\mathbb{E}\,|\varepsilon_t|^{\bar{q}} < \infty$ with $\bar{q} > (d - 1/2)^{-1}$ must also hold (item i. Johansen and Nielsen's (2012) Theorem 10). Finally, we apply the delta method to obtain the limiting distribution of the volatility discovery estimates. To show that, define $\alpha_\perp$ and $\widehat{\alpha}_\perp$ as $\alpha_\perp = \frac{1}{\varsigma}\left[\iota_S - \alpha\left(\alpha'\alpha\right)^{-1}\alpha'\iota_S\right]$ with $\varsigma = \iota_S'\iota_S - \iota_S'\alpha(\alpha'\alpha)^{-1}\alpha'\iota_S$ and $\widehat{\alpha}_\perp = \frac{1}{\widehat{\varsigma}}\left[\iota_S - \widehat{\alpha}\left(\widehat{\alpha}'\widehat{\alpha}\right)^{-1}\widehat{\alpha}'\iota_S\right]$ with $\widehat{\varsigma} = \iota_S'\iota_S - \iota_S'\widehat{\alpha}(\widehat{\alpha}'\widehat{\alpha})^{-1}\widehat{\alpha}'\iota_S$, respectively, such that $\alpha'\alpha_\perp = 0$ and $\widehat{\alpha}'\widehat{\alpha}_\perp = 0$ hold and ι_S denotes a S-dimensional vector of ones. Also, let $\triangledown\alpha_\perp(\alpha) = \frac{\partial\alpha_\perp}{\partial \text{vec}(\alpha)'}$ and $\triangledown\widehat{\alpha}_\perp(\widehat{\alpha}) = \frac{\partial\widehat{\alpha}_\perp}{\partial \text{vec}(\widehat{\alpha})'}$. It then follows that the asymptotic distribution of $\widehat{\alpha}_\perp$ reads

$$\sqrt{T}\left(\widehat{\alpha}_\perp - \alpha_\perp\right) \xrightarrow{d} \mathcal{N}\left(0, \triangledown\alpha_\perp(\alpha)\mathbb{J}V^{-1}\mathbb{J}'\triangledown\alpha_\perp(\alpha)'\right),$$

which finalizes the proof of Theorem 2.

S.2 Price Discovery Overview

The price discovery literature typically adopts the VEC models to approximate the dynamics of prices from a same asset that is traded at multiple exchanges and hence identify the latent efficient price.

Let t_i, $i = 1, 2, \ldots, n$, denote an intraday observation within a trading day t, n be the total number of intraday observations, such that $t - 1 = t_0 < t_1 \ldots < t_n = t$, and T be the total number of trading days, such that $t = 1, \ldots, T$. The general VEC model then reads

$$\Delta P_{t_i} = \gamma\delta' P_{t_{i-1}} + \sum_{j=1}^{q} \Upsilon_j \Delta P_{t_{i-j}} + e_{t_i}, \quad i = 1, 2, ..., n, \tag{S.15}$$

where P_{t_i} is a $S \times 1$ vector that collects the log prices on the S different trading venues, γ and δ are the speed of adjustment coefficients and the cointegrating vector, respectively, and e_{t_i} is a sequence of uncorrelated innovations. Because the market prices share the same stochastic trend (the efficient price), these prices must cointegrate. This implies that γ and δ are $S \times S-1$ matrices with full column rank, such that $\mathrm{rk}\left(\gamma\delta'\right) = S-1$. Furthermore, the cointegrating vector δ is assumed to be $\delta = \left(I_{S-1}, -\iota_{S-1}\right)'$, where ι_{S-1} denotes a vector of ones. In that, price differentials across markets are stationary processes. Based on the approximation in (S.15), we can recover the stochastic trend and ultimately investigate how it is related to the market prices.

Identification of the efficient price arises from one of many types of multivariate random walk decompositions. They decompose observed prices into two components: the $I\left(1\right)$ efficient price and an $I\left(0\right)$ process that is associated with the portion of information that has no permanent impact on prices. Among the several different decompositions, the Granger representation theorem has the advantage of ensuring that the common efficient price is a martingale (Hansen and Lunde, 2006),

$$P_{t_i} = \delta_{\perp}\left[\gamma_{\perp}'\left(I - \sum_{j=1}^{q}\Upsilon_j\right)\delta_{\perp}'\right]^{-1}\gamma_{\perp}'\sum_{j=1}^{i}e_{t_j} + \sum_{j=0}^{\infty}\Xi_j e_{t_{i-j}} + P_{t_0}, \quad i = 1, 2, ..., n, \qquad \text{(S.16)}$$

where $\delta_{\perp} = \iota_S = (1, 1, ..., 1)'$ in the price discovery context, and $\gamma_{\perp}$ is the orthogonal complement of the speed of adjustment parameter that satisfies $\gamma'\gamma_{\perp} = 0$ and $\gamma_{\perp}'\iota_S = 1$. Because the first term of (S.16) is seen as the common stochastic trend, the elements of $\gamma_{\perp}$ are seen as the weights at which market innovations affect the efficient price. In the price discovery context, $\gamma_{\perp}$ is usually referred as the component share (CS) measure. The CS measure has been widely applied to price discovery analysis (Booth et al., 1999, Chu et al., 1999, Harris et al., 2002, Figuerola-Ferretti and Gonzalo, 2010, Scherrer, 2021 among others), where the market associated with the highest element of $\gamma_{\perp}$ is the most important in the price discovery process (see de Jong, 2002 for a precise discussion of $\gamma_{\perp}$ and its relation with other price discovery measures). Finally, Dias et al. (2020) shows that $\gamma_{\perp}$ is invariant to the sampling frequency. This means that the orthogonal complement of γ is constant across prices sampled at alternative sampling intervals.

Another popular price discovery measure is Hasbrouck's (1995) information share (IS) measure. The IS measure gauges the contribution of each market/venue to the unconditional variance of the common stochastic trend (see, for instance, de Jong, 2002, Baillie et al., 2002, and Yan and Zivot, 2010 for an overview and discussion about the properties of the IS and CS measures). More specifically, the IS reads

$$IS_s = \frac{[\psi C_\delta]_s^2}{\psi \Sigma_\delta \psi'}, \ \ s = 1, \dots, S \tag{S.17}$$

where $\Sigma_\delta = \operatorname{Var}(e_{t_i})$, C_δ is such that $\Sigma_\delta = C_\delta C_\delta'$, ψ is the common row of $\delta_\perp \left[\gamma_\perp' \left(I - \sum_{j=1}^{q} \Upsilon_j\right) \delta_\perp'\right]^{-1} \gamma_\perp'$ which follows directly from the fact that $\delta_\perp = \iota_S$, and [.] denotes the sth element of a vector. Finally, Dias et al. (2020) derive a continuous time counterpart of the IS measure in (S.17) and show that the IS measure is not invariant to the sampling frequency, as it converges to $1/S$ as the sampling interval increases. This means that the interpretation of the IS measure should take into account the sapling interval in which prices are sampled.

S.3 Simulation

In terms of assessing the validity of the volatility discovery framework, we analyse a simulation study that is designed to be driven by the two information channels: the price and volatility discovery mechanisms. To exemplify in a pricing model how markets can be affected by two sources of information, we design an example in which markets adjust to the efficient price and volatility. Therefore, there are two channels of information flow: the usual price discovery process and the novel volatility discovery mechanism. Consider a homogeneous asset that is traded in two markets; hence, these prices cointegrate and share the efficient price. We parametrize these prices in such a way they satisfy Assumption 1, i.e., they are continuous-time diffusion processes with time-varying stochastic volatilities and reduced rank drift. As for the data-generating process of the stochastic volatilities, the obvious choice would be to simulate the IVs from a FCVAR model. However, this strategy would add little value to our analysis, because both the RMs and the MLE-based volatility discovery measures are consistent estimates of the market-specific IVs and $\alpha_\perp$, respectively. More interestingly, we investigate the ability of the volatility discovery framework to approximate the dynamics of RMs computed from prices generated from a general stochastic volatility model. That is, we parametrize the data-generating process such that the volatility discovery dynamics are embedded in the spot volatilities. To this extent, we follow the standard practice in the RM literature and model the spot stochastic volatilities as a stationary OrnsteinUhlenbeck (OU) process. The continuous-time price process reads

$$\mathrm{d}p(t) = \begin{bmatrix} 0 & 0 \\ -\gamma_2 & \gamma_2 \end{bmatrix} \Big(\mu_p - p(t)\Big)\mathrm{d}t + \begin{bmatrix} \sigma_1(t) & 0 \\ 0 & \sigma_2(t) \end{bmatrix} \mathrm{d}W(t) \tag{S.18}$$

$$\mathrm{d}V(t) = \begin{bmatrix} -\theta_1 & \theta_1 \\ 0 & -\theta_2 \end{bmatrix} \Big(\mu_V - V(t)\Big)\mathrm{d}t + C\mathrm{d}B(t), \tag{S.19}$$

where $p_t = (p_{1,t}, p_{2,t})'$ and $V(t) = (V_1(t), V_2(t))'$ are 2×1 vectors containing the observed log prices and the logarithm of the instantaneous stochastic volatilities, respectively; μ_p and μ_V are 2×1 vectors of mean parameters; C is 2×2 matrix such that the instantaneous covariance matrix of the log-volatility process is given by $\Lambda = CC'$; $\sigma_s(t) = \exp[\varphi_0 + \varphi_1 V_s(t)]$ with $s = 1, 2$; and W and B are 2×1 vectors of Brownian motions with $\mathrm{Corr}\left(\mathrm{d}W_s(t)\mathrm{d}B_\ell(t)\right) = \rho$ for $s, \ell = 1, 2$. The parameter restrictions in the error correction model are imposed such that the equilibrium relationship between the stochastic volatilities in the two markets is given by $\mathbb{E}(V_1(t)) = \mathbb{E}(V_2(t))$. Without loss of generality, the off-diagonal elements of the spot volatility matrix in (S.18) are set to zero.

The first information channel emerges from the cointegration relationship among these prices. We follow Hasbrouck (1995) and say that market one is the single driver in the price discovery process. In other words, while we impose that the efficient price is exogenous, i.e., the equivalent of a random walk process in discrete time, we allow market one to directly track the efficient price, while market two only learns about changes in the efficient price from market one (market two does not Granger-cause market one).

The second information channel consists on how the market-specific volatilities learn about changes in the efficient volatility (the volatility discovery mechanism). We assume that the stochastic volatilities in the two markets are tied in a long-run equilibrium and hence do not diverge; the efficient and market-specific volatilities are generated by a stationary error correction model in continuous time. For ease of interpretation, we assign market two as the single driver in the volatility discovery process.[2] In turn, while investors in market one impound all relevant information to the efficient price, market two is exclusively responsible for determining the efficient stochastic volatility process.

We simulate prices from the joint continuous time model for price and volatility discovery in (S.18) and (S.19) for $T = 1,700$ trading days (approximately 7 years). Specifically, each day is simulated over the unit interval $t \in [0, 1]$. We normalize 10 seconds to be 1/2,340 so that the interval $[0, 1]$ contains 6.5 hours. In turn, we discretize $[0, 1]$ into $n = 2,340$ intervals with size $\delta = 1/2,340$. The discretized prices are thus comparable to observed prices sampled at one observation every 10 seconds. While the bivariate price process is simulated via the Euler scheme, $V(t)$ is obtained using the exact discretization of the OU process. It follows that discrete prices and stochastic volatilities are obtained using the

[2]The restrictions on a bivariate stationary error correction model in continuous time are less obvious than in its discrete-time counterpart. Specifically, we use the exact discretization of the OU process to map the restrictions implied by the volatility discovery mechanism from discrete time to continuous time.

following iterative scheme:

$$p_{t_{i+1}} = p_{t_i} + \begin{bmatrix} 0 & 0 \\ -\gamma_2 & \gamma_2 \end{bmatrix} \Big(\mu_p - p_{t_i}\Big)\delta + \begin{bmatrix} \sigma_{1,t_i} & 0 \\ 0 & \sigma_{2,t_i} \end{bmatrix} \sqrt{\delta}\,\epsilon^{W}_{t_{i+1}}, \tag{S.20}$$

$$\sigma_{s,t_{i+1}} = \exp\big[\varphi_0 + \varphi_1 V_{s,t_{i+1}}\big], \qquad s = 1, 2 \tag{S.21}$$

$$V_{t_{i+1}} = \mu^*_V + \operatorname{expm}(\Theta\delta)\, V_{t_i} + C_\delta \epsilon^B_{t_{i+1}}, \tag{S.22}$$

where $\Theta = \begin{bmatrix} -\theta_1 & \theta_1 \\ 0 & -\theta_2 \end{bmatrix}$, $\mu^*_V = (I_2 - \operatorname{expm}(\Theta\delta))\,\mu_V$,

$$\begin{pmatrix} \epsilon^W_{t_i} \\ \epsilon^B_{t_{i+1}} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \bullet & \bullet & \bullet \\ 0 & 1 & \bullet & \bullet \\ \rho & \rho & 1 & \bullet \\ \rho & \rho & 0 & 1 \end{pmatrix}\right), \tag{S.23}$$

and $\Lambda_\delta = C_\delta C'_\delta$ with $\Lambda_\delta = \int_0^\delta \operatorname{expm}(u\Theta)\,\Lambda\,\operatorname{expm}(u\Theta')\,du$.

Finally, we choose the parameters in (S.18) and (S.19) to be in accordance with the RM literature (Huang and Tauchen, 2005, Barndorff-Nielsen et al., 2008, Barndorff-Nielsen et al., 2011, among others), as the stationary error correction model in (S.19) is a special case of the single-factor log-linear stochastic volatility model. We set the free parameters in (S.18) and (S.19) as follows:

$$\begin{pmatrix} \mu_p \\ \gamma_2 \\ \varphi_0 \\ \varphi_1 \\ \mu_V \\ \theta_1 \\ \theta_2 \\ \Lambda_{1,1} \\ \Lambda_{1,2} \\ \Lambda_{2,2} \\ \rho \end{pmatrix} = \begin{pmatrix} 0.003 \\ 223 \\ 0 \\ 0.125 \\ 0 \\ 0.150 \\ 0.025 \\ 1 \\ 0 \\ 1 \\ -0.30 \end{pmatrix}, \tag{S.24}$$

where $\Lambda_{1,1}$, $\Lambda_{1,2}$ and $\Lambda_{2,2}$ are the elements of Λ. Notably, values assigned to μ_p, μ_V, φ_1, θ_2 Λ and ρ follow from Barndorff-Nielsen et al. (2008) and Barndorff-Nielsen et al. (2011), while $\varphi_0 = 0$ follows from Huang and Tauchen (2005). We choose θ_1 so that the volatility process in the first market (satellite market) is less persistent than the volatility process associated with the leading market. It follows that because the satellite market does not contribute to the volatility discovery, a less persistent stochastic volatility process increases the speed of adjustment to the efficient stochastic volatility given by the second market. For each replication, VEC and FCVAR models are fitted to intraday prices and daily realized variances, respectively, so that price and volatility discovery measures are recorded. We report the mean and standard deviations computed across 1,000 replications.

The results confirm that $\alpha_\perp$ is a valid measure and successfully identifies the volatility discovery mechanism, whereas $\gamma_\perp$ is able to capture only information from the price discovery channel (see Tables S.1 and S.2). Specifically, we first confirm that the FCVAR model approximates well the dynamics of the RMs and thus opens room to assess the validity of the volatility discovery framework. We find that the RMs are characterized by nonstationary and mean reverting long memory, which is in line with the usual hyperbolic decay of the empirical autocorrelation function associated with RMs. Secondly, we document that the FCVAR model is able to correctly identify the structural features associated with the volatility discovery dynamics. The estimates of the speed of adjustment parameters verify the Granger causality structure in (S.19), i.e., α_1 is negative and significantly different from zero (RM in market one adjusts to changes in the RM of market two), whereas α_2 is not statistically different from zero, implying that the RM in the first market does not Granger-cause the RM in market two. Estimates of α are also fairly stable across all replications. Therefore, the use of the FCVAR model to approximate the dynamics of the RMs and the use of the orthogonal complement of the speed of adjustment parameters to quantify how the different markets impound information to the fractional common stochastic trend (efficient volatility) enable inference about the volatility discovery mechanism.

S.4 Auxiliary Results

This section presents additional empirical results that reinforce the main findings discussed in the paper. Table S.3 provides the data details, while Table S.4 reports the lag order of the FCVAR model and the associated p-values for the heteroskedastic robust LM test for serial correlation in the residuals. The results of the residual analysis can be found in Table S.5, and the estimates of α are presented in Table S.6. Table S.7 displays the p-values for the null hypothesis $\alpha_{\perp,s} = 0$ with $s \in (\mathrm{T}, \mathrm{P}, \mathrm{N})$, providing the necessary information to implement the Holm-Bonferroni procedure. Moreover, Table S.8 offers probit regressions that explore the relationship between quoting intensity and price and volatility discovery. Specifically, we regress a binary variable that takes the value of 1 if a given market leads the price (volatility) discovery process and 0 otherwise on quote intensity ratio defined as the ratio of the daily average of quotes in the first market over the sum of daily averages from both markets, volatility (price) discovery measures and a dummy variable that returns 1 if the stock is listed in a given market. In line with our previous explanation, this exercise shows that price discovery is more associated with quoting activity than volatility discovery is, as the probability of leading the price discovery process is positively (statistically significant) related with quote intensity and the volatility discovery measure. Finally, leadership in the volatility discovery process appears to cause price discovery, which highlights the relevant informational content identified by our framework.

S.5 Auxiliary Tables

Table S.1: **Simulation Results: Example - Price Discovery**

We report the mean and standard deviations (in brackets) of the estimates of the VEC parameters (γ and δ) and the price discovery measures ($\gamma_\perp$) and IS computed across 1,000 replications from models simulated using the data generation process in our price model in (S.18) and (S.19). For each replication, intraday prices and stochastic volatilities are simulated via an Euler scheme over the unit interval $t \in [0,1]$ with steps of size 1/2,340 which corresponds to 10 second frequency. In turn, the interval $t \in [0,1]$ contains 6.5 hours. We simulate 1,700 days (about 7 years). Because (S.18) is a continuous-time reduced-rank diffusion process with time-varying stochastic volatilities, the Euler discretization yields true parameters that are comparable with the estimates of the discrete VEC model. It follows that $\gamma = (0, 0.10)'$, $\delta = (1, -1)'$ and $\gamma_\perp = (1, 0)'$.

VEC approximation of high-frequency log-prices

$$\Delta\left(\begin{bmatrix} p_{1,t_i} \\ p_{2,t_i} \end{bmatrix}\right) = \underbrace{\begin{bmatrix} \underset{(1.57\times 10^4)}{-0.00} \\ \underset{(1.54\times 10^4)}{0.10} \end{bmatrix}}_{\widehat{\gamma}_\delta} \underbrace{\begin{bmatrix} 1 & -1 \end{bmatrix}}_{\delta'} \begin{bmatrix} p_{1,t_{i-1}} \\ p_{2,t_{i-1}} \end{bmatrix} + \begin{bmatrix} e_{1,t_i} \\ e_{2,t_i} \end{bmatrix}$$

Price discovery measures

$$\widehat{\gamma}_\perp = \left(\underset{(16\times 10^4)}{1.00}, \quad \underset{(16\times 10^4)}{0.00}\right)'$$

$$IS = \left(\underset{(1.82\times 10^4)}{1.00}, \quad \underset{(1.82\times 10^4)}{0.00}\right)'$$

Table S.2: **Simulation Results: Example - Volatility Discovery**

We report the mean and standard deviations (in brackets) of the estimates of the FCVAR parameters (d, b, α and β) and the volatility discovery measures ($\alpha_\perp$) computed across 1,000 replications from models simulated using the data generation process in our price model in (S.18) and (S.19). For each replication, intraday prices and stochastic volatilities are simulated via an Euler scheme over the unit interval $t \in [0,1]$ with steps of size 1/2,340 which corresponds to 10 second frequency. In turn, the interval $t \in [0,1]$ contains 6.5 hours. We simulate 1,700 days (about 7 years) and the daily realized measures are computed using the realized variance estimator defined as the sum of the squared intraday returns.

FCVAR approximation of daily realized measures

$$\Delta^{\underset{(0.05)}{0.88}}\left(\begin{bmatrix} R_{1,t_i} \\ R_{2,t_i} \end{bmatrix}\right) = \underbrace{\begin{bmatrix} \underset{(0.01)}{-0.11} \\ \underset{(0.01)}{0.01} \end{bmatrix}}_{\widehat{\alpha}} \underbrace{\begin{bmatrix} 1 & \underset{(0.05)}{-1.04} \end{bmatrix}}_{\widehat{\beta}'} \begin{bmatrix} R_{1,t_{i-1}} \\ R_{2,t_{i-1}} \end{bmatrix} + \sum_{j=1}^{\kappa} \widehat{\Gamma}_j \Delta^{\underset{(0.05)}{0.88}} L^j_{\underset{(0.05)}{0.88}} R_t + \begin{bmatrix} \varepsilon_{1,t_i} \\ \varepsilon_{2,t_i} \end{bmatrix}$$

Volatility discovery measures

$$\widehat{\alpha}_\perp = \left(\underset{(0.09)}{0.01}, \quad \underset{(0.09)}{0.99}\right)'$$

Table S.3: **Data Description**

We report summary statistics for raw and cleaned data considering Nasdaq, Arca, and Nyse. The first three columns present the number of quotes (in millions) for each stock before any cleaning filter (raw data). The following three columns (Clean obs) display the total number of quotes (in millions) after the implementation of the cleaning procedure. The following three columns (Avg obs per day) stand for the daily average (in thousands) of quotes. The last three columns report the total number of days we have for each stock for the time span 01/01/2007 to 31/12/2013.

	Initial Obs (Million)			Clean Obs (Million)			Avg obs per day (Thousand)			Number of days		
	Nasdaq	Arca	Nyse	Nasdaq	Arca	Nyse	Nasdaq	Arca	Nyse	Nasdaq	Arca	Nyse
AA	237	138	247	25	23	27	14.32	13.27	15.14	1735	1762	1762
AAPL	459	256	-	31	30	-	17.50	17.24	-	1762	1762	-
BAC	523	292	503	31	30	34	17.85	17.28	19.05	1735	1762	1762
BRKB	98	68	-	11	10	-	9.12	8.35	-	1183	1142	-
C	-	319	549	-	30	32	-	16.93	18.39	-	1762	1762
CSCO	210	67	-	14	11	-	17.72	14.79	-	777	777	-
F	137	70	172	13	12	15	13.24	11.77	14.91	982	1009	1009
GE	363	214	427	29	27	31	16.53	15.58	17.78	1735	1762	1762
GM	202	90	174	16	15	19	11.81	10.83	13.34	1364	1391	1391
GOOG	149	124	-	19	19	-	10.77	10.79	-	1762	1762	-
HPQ	326	167	277	26	24	28	14.84	13.67	15.86	1735	1762	1762
IBM	122	102	149	21	20	25	11.96	11.50	14.13	1735	1762	1762
JCP	175	113	149	20	20	22	11.55	11.14	12.26	1735	1762	1762
JNJ	280	137	251	25	22	27	18.37	15.84	19.76	1364	1391	1391
JPM	696	345	542	32	31	33	18.43	17.52	18.64	1735	1762	1762
KO	244	123	205	23	21	25	13.27	11.90	14.44	1735	1762	1762
MO	178	95	204	22	19	26	12.40	10.83	14.90	1735	1762	1762
MRK	271	151	244	25	23	27	14.19	13.01	15.47	1735	1762	1762
MRVL	252	101	-	24	19	-	13.46	11.03	-	1762	1762	-
MS	233	416	-	28	27	-	16.03	15.40	-	1735	1762	-
MSFT	669	228	-	33	28	-	18.59	16.13	-	1762	1762	-
NOK	199	120	187	22	21	23	12.46	11.70	12.99	1735	1762	1762
ORCL	494	181	-	30	26	-	17.15	14.49	-	1762	1762	-
PFE	309	159	342	27	25	30	15.29	13.92	17.12	1735	1762	1762
PG	283	136	198	25	22	26	14.41	12.54	14.77	1735	1762	1762
VZ	264	141	257	25	23	29	14.44	13.03	16.19	1735	1762	1762
WFC	-	305	427	-	29	31	-	16.67	17.81	-	1762	1762
WMT	269	144	251	25	23	28	14.40	12.98	16.01	1735	1762	1762
XOM	503	337	417	31	31	33	18.10	17.56	18.84	1735	1762	1762
YHOO	367	121	-	26	22	-	14.87	12.27	-	1762	1761	-

Table S.4: **Model Specification: Lag Length Selection**

We report results for 30 stocks considering the Nasdaq and Arca, Nasdaq and Nyse, Arca and Nyse, and Nasdaq, Arca and Nyse market combinations. For each set of results, κ accounts for the selected lag length in the FCVAR model; and LM_s $s \in (T, P, N)$ brings the p-values associated with the heteroskedastic robust LM test for serial correlation at lag 10 for the residuals in sth equation in the FCVAR model, respectively. The null hypothesis of the LM test is that the process is serially uncorrelated. "-" implies that the stock is not traded in at least one of the two trading venues during the entire sample. Finally, we set $\kappa = 0$ for the NOK asset in both Nasdaq-Nyse and Arca-Nyse systems because the roots of the characteristic polynomial lie inside the transformed unit circle for any $\kappa > 0$.

	Nasdaq-Arca			Nasdaq-Nyse			Arca-Nyse			Nasdaq-Arca-Nyse			
	κ	LM_T	LM_P	κ	LM_T	LM_N	κ	LM_P	LM_N	κ	LM_T	LM_P	LM_N
AA	0	0.87	0.83	0	0.90	0.97	0	0.76	0.94	0	0.90	0.86	0.96
AAPL	0	0.18	0.19	-	-	-	-	-	-	-	-	-	-
BAC	0	0.48	0.48	0	0.48	0.56	0	0.63	0.69	0	0.47	0.48	0.55
BRKB	0	0.29	0.23	1	0.10	0.43	2	0.61	0.49	-	-	-	-
C	-	-	-	-	-	-	0	0.21	0.30	-	-	-	-
CSCO	0	0.33	0.18	-	-	-	-	-	-	-	-	-	-
F	0	0.29	0.15	0	0.24	0.30	0	0.23	0.45	0	0.37	0.22	0.41
GE	0	0.23	0.19	0	0.24	0.29	0	0.18	0.25	0	0.22	0.18	0.27
GM	0	0.19	0.12	0	0.20	0.10	0	0.24	0.18	0	0.19	0.13	0.10
GOOG	0	0.01	0.01	-	-	-	-	-	-	-	-	-	-
HPQ	0	0.16	0.10	0	0.15	0.19	0	0.10	0.23	0	0.20	0.12	0.22
IBM	1	0.15	0.12	1	0.13	0.13	2	0.20	0.20	0	0.03	0.02	0.04
JCP	0	0.43	0.27	0	0.44	0.39	0	0.22	0.36	0	0.34	0.19	0.32
JNJ	0	0.71	0.56	0	0.77	0.68	0	0.75	0.72	0	0.79	0.67	0.69
JPM	0	0.91	0.82	0	0.91	0.84	0	0.80	0.81	0	0.92	0.83	0.85
KO	3	0.31	0.14	0	0.14	0.24	0	0.02	0.15	0	0.14	0.05	0.26
MO	0	0.59	0.30	0	0.44	0.42	0	0.25	0.34	0	0.56	0.28	0.44
MRK	0	0.84	0.57	0	0.77	0.85	0	0.57	0.87	0	0.84	0.56	0.89
MRVL	0	0.30	0.23	-	-	-	-	-	-	-	-	-	-
MS	0	0.41	0.31	-	-	-	-	-	-	-	-	-	-
MSFT	0	0.44	0.38	-	-	-	-	-	-	-	-	-	-
NOK	2	0.25	0.18	0	0.06	0.03	0	0.03	0.03	3	0.09	0.06	0.06
ORCL	0	0.33	0.32	-	-	-	-	-	-	-	-	-	-
PFE	0	0.17	0.14	0	0.28	0.26	0	0.35	0.25	0	0.19	0.17	0.17
PG	0	0.21	0.15	0	0.22	0.23	0	0.17	0.25	0	0.24	0.16	0.26
VZ	0	0.37	0.35	0	0.41	0.42	0	0.39	0.43	0	0.41	0.39	0.42
WFC	0	0.43	0.40	-	-	-	-	-	-	-	-	-	-
WMT	0	0.30	0.16	0	0.27	0.28	0	0.18	0.29	-	-	-	-
XOM	3	0.16	0.16	3	0.12	0.16	3	0.09	0.10	0	0.01	0.01	0.01
YHOO	0	0.73	0.64	-	-	-	-	-	-	-	-	-	-

Table S.5: **Model Specification: Residuals Analysis**

We report the residuals analysis for 30 stocks considering the Nasdaq and Arca, Nasdaq and Nyse, Arca and Nyse, and Nasdaq, Arca and Nyse market combinations. For each set of results, LM_s $s \in (T, P, N)$ brings the p-values associated with the heteroskedastic robust LM test for serial correlation at lag 15 for the residuals in sth equation in the FCVAR model. The null hypothesis of the LM test is that the process is serially uncorrelated. "-" implies that the stock is not traded in at least one of the two trading venues during the entire sample.

	Nasdaq-Arca		Nasdaq-Nyse		Arca-Nyse		Nasdaq-Arca-Nyse		
	LM_T	LM_P	LM_T	LM_N	LM_P	LM_N	LM_T	LM_P	LM_N
AA	0.80	0.73	0.84	0.95	0.65	0.91	0.90	0.86	0.96
AAPL	0.03	0.01	-	-	-	-	-	-	-
BAC	0.47	0.47	0.49	0.56	0.63	0.71	0.47	0.48	0.55
BRKB	0.26	0.25	-	-	-	-	-	-	-
C	-	-	-	-	0.22	0.30	-	-	-
CSCO	0.28	0.14	-	-	-	-	-	-	-
F	0.26	0.13	0.20	0.27	0.20	0.42	0.37	0.22	0.41
GE	0.19	0.15	0.20	0.24	0.13	0.19	0.22	0.18	0.27
GM	0.17	0.11	0.18	0.09	0.21	0.16	0.19	0.13	0.10
GOOG	0.01	0.01	-	-	-	-	-	-	-
HPQ	0.16	0.10	0.15	0.19	0.10	0.23	0.20	0.12	0.22
IBM	0.15	0.13	0.13	0.12	0.21	0.20	0.03	0.02	0.04
JCP	0.33	0.20	0.39	0.30	0.20	0.28	0.34	0.19	0.32
JNJ	0.64	0.47	0.70	0.59	0.62	0.61	0.79	0.67	0.69
JPM	0.90	0.81	0.90	0.84	0.78	0.81	0.92	0.83	0.85
KO	0.33	0.16	0.10	0.18	0.01	0.11	0.14	0.05	0.26
MO	0.48	0.19	0.31	0.33	0.12	0.27	0.56	0.28	0.44
MRK	0.51	0.25	0.49	0.61	0.24	0.59	0.84	0.56	0.89
MRVL	0.11	0.07	-	-	-	-	-	-	-
MS	0.39	0.30	-	-	-	-	-	-	-
MSFT	0.46	0.46	-	-	-	-	-	-	-
NOK	0.22	0.15	0.04	0.02	0.02	0.02	0.09	0.06	0.06
ORCL	0.24	0.22	-	-	-	-	-	-	-
PFE	0.12	0.09	0.13	0.12	0.12	0.14	0.19	0.17	0.17
PG	0.15	0.10	0.16	0.18	0.11	0.19	0.24	0.16	0.26
VZ	0.28	0.25	0.30	0.33	0.27	0.36	0.41	0.39	0.42
WFC	-	-	-	-	0.51	0.42	-	-	-
WMT	0.11	0.05	0.10	0.12	0.06	0.13	0.30	0.16	0.31
XOM	0.12	0.13	0.10	0.12	0.07	0.07	0.01	0.01	0.01
YHOO	0.58	0.47	-	-	-	-	-	-	-

Table S.6: **Speed of Adjustment**

We report the speed of adjustments coefficients for 30 assets considering the Nasdaq and Arca, Nasdaq and Nyse, Arca and Nyse, and Nasdaq, Arca and Nyse market combinations. The FCVAR parameters are computed using the MLE estimator of Johansen and Nielsen (2012) where $\mathrm{rk}\,(\alpha\beta') = S-1$ with S denoting the number of markets, κ is chosen according to Table S.4, and $d=b$. For each set of of market combinations, we report the estimates of the elements of the $S(S-1)\times 1$ vector $\mathrm{vec}(\alpha)$ and their one-sided significance levels . The symbols *, ** and *** denote rejection at the 10%, 5% and 1% levels of the null hypothesis of $\widehat{\alpha}_i' = 0$, for $i = 1,\ldots,S(S-1)$. "-" implies that the stock is not traded in at least one of the two trading venues during the entire sample.

	Nasdaq-Arca		Nasdaq-Nyse		Arca-Nyse		Nasdaq -Arca-Nyse					
	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_{1,1}$	$\widehat{\alpha}_{2,1}$	$\widehat{\alpha}_{3,1}$	$\widehat{\alpha}_{1,2}$	$\widehat{\alpha}_{2,2}$	$\widehat{\alpha}_{3,2}$
AA	-0.42	0.55^{*}	-0.46^{**}	0.38^{*}	-0.43^{**}	0.47^{**}	-0.65	0.33	0.15	0.25	-0.71^{*}	0.31
AAPL	-0.13	0.75^{*}	-	-	-	-	-	-	-	-	-	-
BAC	-0.26	0.73^{*}	-0.10	0.76^{***}	-0.23	0.63^{**}	-0.27	0.73	0.62	0.23	-0.74	0.19
BRKB	-0.57^{**}	0.29	-	-	-	-	-	-	-	-	-	-
C	-	-	-	-	0.19	1.16^{***}	-	-	-	-	-	-
CSCO	-0.32	0.68^{***}	-	-	-	-	-	-	-	-	-	-
F	-1.07^{**}	-0.10	0.08	1.02^{***}	0.40	1.34^{***}	-0.89	0.08	0.06	1.29^{**}	0.32	1.26^{**}
GE	-0.92^{**}	0.00	-0.10	0.66^{**}	0.18	0.83^{***}	-0.91^{*}	0.04	-0.03	1.05^{*}	0.19	0.89^{*}
GM	-1.12^{***}	-0.11	-0.35^{*}	0.62^{***}	-0.07	0.75^{***}	-1.18^{**}	-0.13	-0.05	1.04^{**}	0.08	0.84^{*}
GOOG	-0.57^{**}	0.21	-	-	-	-	-	-	-	-	-	-
HPQ	-1.03^{***}	-0.08	-0.77^{***}	0.05	-0.51^{**}	0.29	-1.20^{***}	-0.26	-0.30	0.60	-0.38	0.47
IBM	-0.53^{*}	0.40	-0.39	0.32	-0.17	0.55^{*}	-0.73^{**}	0.20	0.18	0.31	-0.61^{**}	0.20
JCP	-0.58^{**}	0.41^{*}	-0.73^{***}	0.12	-0.58^{**}	0.19	-0.90^{***}	0.10	0.03	0.32	-0.67^{**}	0.16
JNJ	-0.54^{**}	0.45^{**}	-0.54^{***}	0.22	-0.53^{***}	0.16	-0.75^{***}	0.26	0.17	0.29	-0.68^{***}	0.07
JPM	-0.34	0.65^{**}	-0.15	0.74^{***}	-0.26	0.48^{**}	-0.34	0.66^{*}	0.68^{*}	0.30	-0.64	0.10
KO	-0.65^{*}	0.14	-0.61^{***}	0.21	-0.58^{***}	0.14	-0.80^{***}	0.10	0.09	0.23	-0.63^{**}	0.14
MO	-0.54^{**}	0.42^{**}	-0.60^{***}	0.22	-0.60^{***}	0.18	-0.82^{***}	0.14	0.12	0.31	-0.66^{**}	0.14
MRK	-0.56^{**}	0.45^{**}	-0.91^{***}	0.01	-0.84^{***}	0.03	-1.02^{***}	0.01	-0.04	0.16	-0.83^{***}	0.07
MRVL	-0.78^{**}	0.21	-	-	-	-	-	-	-	-	-	-
MS	-0.35	0.59^{*}	-	-	-	-	-	-	-	-	-	-
MSFT	0.14	1.12^{***}	-	-	-	-	-	-	-	-	-	-
NOK	-1.18^{*}	-0.25	-0.42^{*}	0.57^{**}	-0.37	0.52^{*}	-1.18	-0.33	-0.33	0.82	0.07	0.78
ORCL	-0.07	0.96^{***}	-	-	-	-	-	-	-	-	-	-
PFE	-0.87^{**}	0.13	-0.57^{***}	0.25	-0.51^{**}	0.27	-1.02^{**}	-0.03	-0.09	0.59	-0.41	0.44
PG	-0.29	0.63^{***}	-0.59^{***}	0.35^{**}	-0.63^{***}	0.21	-0.58^{**}	0.39	0.33	-0.02	-0.87^{***}	0.02
VZ	-0.61^{***}	0.40^{**}	-0.63^{***}	0.27^{*}	-0.64^{***}	0.20	-0.83^{***}	0.17	0.18	0.35	-0.68^{**}	0.17
WFC	-	-	-	-	-0.44^{**}	0.39^{**}	-	-	-	-	-	-
WMT	-0.68^{**}	0.29	-0.49^{**}	0.47^{**}	-0.33	0.57^{**}	-0.86^{**}	0.15	0.13	0.54	-0.39	0.51
XOM	-2.45^{***}	-1.52	-0.77^{*}	0.03	0.49	1.12^{**}	-0.98^{**}	0.03	0.01	-0.17	-1.16^{**}	-0.23
YHOO	-1.01^{***}	-0.01	-	-	-	-	-	-	-	-	-	-

Table S.7: p-values Volatility Discovery Measures

We report the p-values of the null hypothesis $\alpha_{\perp,s} = 0$ with $s \in (T, P, N)$ for 30 assets considering the Nasdaq and Arca, Nasdaq and Nyse, Arca and Nyse, and Nasdaq, Arca and Nyse market combinations. T, P, and N denote Nasdaq, Arca and Nyse, respectively. The FCVAR parameters are computed using the MLE estimator of Johansen and Nielsen (2012) where $\mathrm{rk}(\alpha\beta') = 1$, κ is chosen according to Table 1, and $d = b$. "-" implies that the stock is not traded in at least one of the two trading venues during the entire sample.

	Nasdaq-Arca		Nasdaq-Nyse		Arca-Nyse		Nasdaq-Arca-Nyse		
	$\alpha_{\perp,T}$	$\alpha_{\perp,P}$	$\alpha_{\perp,T}$	$\alpha_{\perp,N}$	$\alpha_{\perp,P}$	$\alpha_{\perp,N}$	$\alpha_{\perp,T}$	$\alpha_{\perp,P}$	$\alpha_{\perp,N}$
AA	[0.06]	[0.13]	[0.07]	[0.04]	[0.03]	[0.04]	[0.28]	[0.23]	[0.08]
AAPL	[0.05]	[0.39]	-	-	-	-	-	-	-
BAC	[0.06]	[0.29]	[0.01]	[0.37]	[0.03]	[0.24]	[0.07]	[0.33]	[0.46]
BRKB	[0.12]	[0.01]	-	-	-	-	-	-	-
C	-	-	-	-	[0.00]	[0.70]	-	-	-
CSCO	[0.00]	[0.11]	-	-	-	-	-	-	-
F	[0.56]	[0.04]	[0.01]	[0.58]	[0.00]	[0.84]	[0.45]	[0.02]	[0.82]
GE	[0.50]	[0.03]	[0.01]	[0.37]	[0.00]	[0.72]	[0.45]	[0.03]	[0.76]
GM	[0.60]	[0.01]	[0.01]	[0.09]	[0.00]	[0.40]	[0.61]	[0.03]	[0.44]
GOOG	[0.25]	[0.03]	-	-	-	-	-	-	-
HPQ	[0.60]	[0.00]	[0.43]	[0.00]	[0.15]	[0.03]	[0.77]	[0.14]	[0.01]
IBM	[0.12]	[0.06]	[0.17]	[0.12]	[0.07]	[0.33]	[0.26]	[0.18]	[0.05]
JCP	[0.08]	[0.02]	[0.32]	[0.00]	[0.22]	[0.01]	[0.44]	[0.28]	[0.03]
JNJ	[0.03]	[0.01]	[0.10]	[0.00]	[0.17]	[0.00]	[0.22]	[0.30]	[0.01]
JPM	[0.04]	[0.18]	[0.00]	[0.28]	[0.04]	[0.17]	[0.04]	[0.26]	[0.47]
KO	[0.38]	[0.09]	[0.12]	[0.00]	[0.20]	[0.00]	[0.37]	[0.27]	[0.00]
MO	[0.04]	[0.01]	[0.14]	[0.00]	[0.18]	[0.00]	[0.32]	[0.25]	[0.01]
MRK	[0.04]	[0.01]	[0.47]	[0.00]	[0.44]	[0.00]	[0.56]	[0.39]	[0.00]
MRVL	[0.28]	[0.02]	-	-	-	-	-	-	-
MS	[0.08]	[0.20]	-	-	-	-	-	-	-
MSFT	[0.01]	[0.62]	-	-	-	-	-	-	-
NOK	[0.64]	[0.07]	[0.03]	[0.08]	[0.12]	[0.16]	[0.66]	[0.21]	[0.38]
ORCL	[0.00]	[0.39]	-	-	-	-	-	-	-
PFE	[0.38]	[0.02]	[0.14]	[0.01]	[0.13]	[0.02]	[0.56]	[0.14]	[0.04]
PG	[0.01]	[0.13]	[0.04]	[0.00]	[0.13]	[0.00]	[0.10]	[0.49]	[0.00]
VZ	[0.04]	[0.00]	[0.09]	[0.00]	[0.17]	[0.00]	[0.24]	[0.20]	[0.02]
WFC	-	-	-	-	[0.03]	[0.02]	-	-	-
WMT	[0.19]	[0.02]	[0.04]	[0.03]	[0.01]	[0.10]	[0.34]	[0.07]	[0.19]
XOM	[0.96]	[0.00]	[0.48]	[0.07]	[0.04]	[0.77]	[0.50]	[0.69]	[0.00]
YHOO	[0.51]	[0.01]	-	-	-	-	-	-	-

Table S.8: **Price (Volatility) Discovery Measures and Quoting Intensity**

We report results from probit regressions relating price and volatility discovery measures, Panels A and B, respectively, to quoting intensity and control variables. We consider two market combinations: Nasdaq-Arca and Nyse-Nasdaq. We omit the third market combination, as transitivity would violate the i.i.d. assumption required for the estimation of probit models. The dependent variable in Panel A is an indicator variable that takes the value of 1 if the price discovery measure associated with the first market is greater than 0.5, and 0 otherwise. There are three independent variables: 'Quote Intensity ratio' is the ratio of the daily average of quotes in the first market over the sum of daily averages from both markets; 'Listed Stock' is a dummy variable that takes the value of 1 if the stock is listed in the first market, and 0 otherwise; and 'Volatility Discovery' is the element of $\alpha_\perp$ associated with the the first market. The dependent variable in Panel B is an indicator variable that takes the value of 1 if the volatility discovery measure associated with the first market is greater than 0.5, and 0 otherwise. There are three independent variables: 'Quote Intensity ratio', 'Listed Stock', and 'Price Discovery', that is the element of $\gamma_\perp$ associated with the the first market. The regression in column 1 uses only the first independent variable, 'Quote Intensity ratio', the regression in the second column uses 'Quote Intensity ratio' and 'Listed Stock', and the regression in the third column uses all the independent variables. The robust standard errors are in parentheses below the coefficient estimates and ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Wald-test gives the p-value of the Wald statistic of the joint significance of the parameters associated with the independent variables; Log Likelihood is the value of the log likelihood function attained at the parameter estimates; R^2 is the pseudo R^2 goodness-of-fit-statistic in the context of probit models; and No. of obs. is the total number of observations used in the regression analysis.

Panel A. Price Discovery

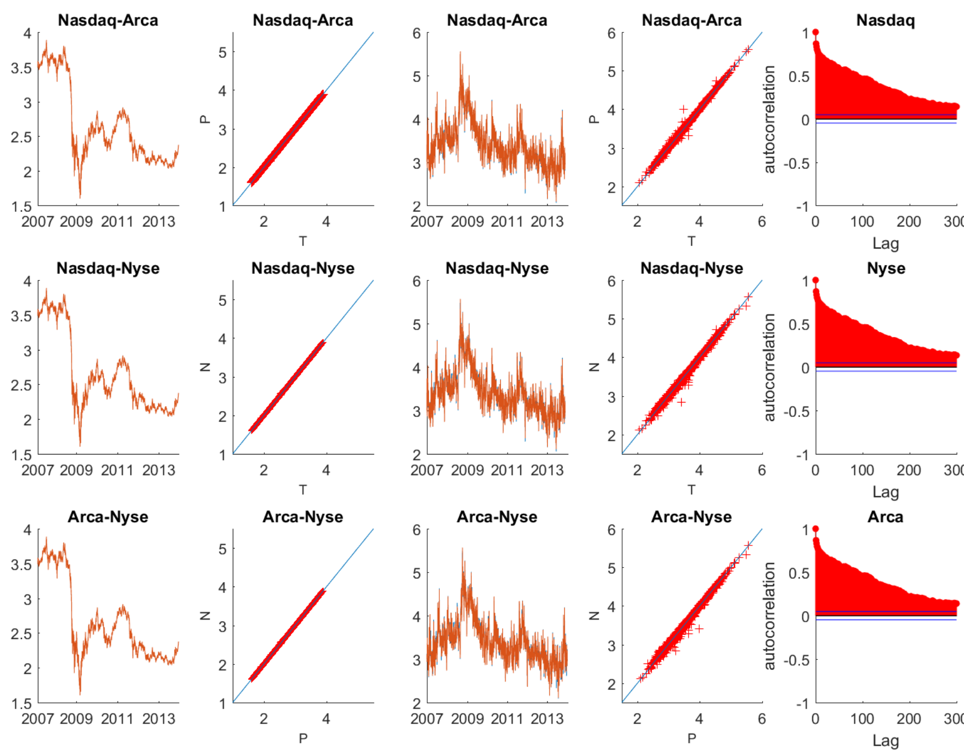
Explanatory Variables	1	2	3
Quote Intensity ratio	6.48^{***}	5.88^{**}	6.20^{**}
	(2.23)	(2.82)	(2.79)
Listed Stock		-0.20	-0.42
		(0.54)	(0.57)
Volatility Discovery: $\alpha_{\perp,1}$			0.75^{*}
			(0.44)
Wald-test	0.00	0.01	0.01
Log Likelihood	-26.39	-26.30	-25.17
pseudo-R^2	0.18	0.18	0.21
No. of obs.	47	47	47

Panel B. Volatility Discovery

Explanatory Variables	1	2	3
Quote Intensity ratio	-3.52^{*}	-1.01	-1.14
	(1.93)	(2.32)	(2.55)
Listed Stock		0.96^{**}	0.96^{**}
		(0.47)	(0.47)
Price Discovery: $\gamma_{\perp,1}$			0.26
			(1.97)
Wald-test	0.07	0.02	0.05
Log Likelihood	-30.60	-28.37	-28.36
pseudo-R^2	0.06	0.13	0.13
No. of obs.	47	47	47

Figure 1: **Prices and Realized Measures at Nasdaq, Arca and Nyse markets: AA (Alcoa)**

The figure consists of three panels with plots of intraday log-prices and daily log-RMs at different market combinations for AA (Alcoa). The first row presents plots for the Nasdaq-Arca market combination, while the second and third panels refer to the Nasdaq-Nyse and Arca-Nyse market combinations. The first column displays plots for the intraday log-prices in the different market combinations. The second column presents scatter plots around the 45 degrees line of intraday log-prices. Specifically, we denote Nasdaq as 'T', Arca as 'P', and Nyse as 'N'. The third column presents plots of daily logarithmic estimates of IV using the realized kernel estimator of Barndorff-Nielsen et al. (2011). The fourth column shows scatter plots around the 45 degrees line of daily logarithmic RMs. Finally, the fifth column displays three plots with sample autocorrelation function of the logarithmic realized kernel estimates for Nasdaq, Nyse, and Arca.



References

Baillie, R. T., G. G. Booth, Y. Tse, and T. Zabotina (2002). Price discovery and common factor models. *Journal of Financial Markets 5*, 309–321.

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

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2011). Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics 162*, 149 –169.

Booth, G. G., R. W. So, and Y. Tseh (1999). Price discovery in the German equity index derivatives markets. *The Journal of Futures Markets 19*, 619–643.

Chu, Q. C., W. G. Hsieh, and Y. Tse (1999). Price discovery on the S&P 500 index markets: An analysis of spot index, index futures and spdrs. *International Review of Financial Analysis 8*, 21–34.

Davidson, J. (1994). *Stochastic Limit Theory. An Introduction for Econometricians.* Oxford University Press.

de Jong, F. (2002). Measures of contributions to price discovery: a comparison. *Journal of Financial Markets 5*, 323–327.

Dias, G. F., M. Fernandes, and C. M. Scherrer (2020, 01). Price discovery in a continuous-time setting. *Journal of Financial Econometrics*.

Figuerola-Ferretti, I. and J. Gonzalo (2010). Modelling and measuring price discovery in commodity markets. *Journal of Econometrics 158*(1), 95–107.

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics 24, N0. 2*, 127–161.

Harris, F. H., T. H. McInish, and R. A. Wood (2002). Security price adjustment across exchanges: an investigation of common factor components for dow stocks. *Journal of Financial Markets 5*, 277–308.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The Journal of Finance 50*, 1175–1198.

Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics 3*, 456–499.

Johansen, S. and M. Ø. Nielsen (2012). Likelihood inference for a fractionally cointegrated vector autoregressive model. *Econometrica 80*(6), 2667–2732.

Kessler, M. and A. Rahbek (2001). Asymptotic likelihood based inference for co-integrated homogenous gaussian diffusions. *Scandinavian Journal of Statistics 28*, 455–470.

Newey, W. K. and D. McFadden (1994, January). Large sample estimation and hypothesis testing. In R. F. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume 4 of *Handbook of Econometrics*, Chapter 36, pp. 2111–2245. Elsevier.

Scherrer, C. M. (2021). Information processing on equity prices and exchange rate for cross-listed stocks. *Journal of Financial Markets 54*.

Stockmarr, A. and M. Jacobsen (1994). Gaussian diffusions and autoregressive processes: Weak convergence and statisticalinference. *Scandinavian Journal of Statistics 21*, 403–419.

Yan, B. and E. Zivot (2010). A structural analysis of price discovery measures. *Journal of Financial Markets 13*, 1–19.