CEIS Tor Vergata Research Papers

ISSN 2610-931X

CEIS Tor Vergata

RESEARCH PAPER SERIES

Vol. 21, Issue 6, No. 568 – November 2023

Funding Liquidity and Stocks' Market Liquidity: Structural Estimation from High-Frequency Data

Gian Piero Aielli and Davide Pirino

Funding Liquidity and Stocks' Market Liquidity: Structural Estimation From High-Frequency Data*

Gian Piero Aielli† Davide Pirino‡

November 2, 2023

Abstract

In accordance with trade signals that operate in the market, we design a micro-founded structural model of price formation that features partially informed and noise traders. The former only have information on whether a trend in the latent price dynamic is underway. Without any trend, the partially informed agents do not trade, and prices do not update unless a noise agent activates. Assuming market efficiency, we impose zero expected net profit per trade. With dedicated parametric assumptions, we analytically derive the model's likelihood, which allows reliable daily estimates (exclusively based on intra-day transaction prices) of the stocks' market liquidities and funding liquidity (and their estimation errors).

Theory predicates that stocks' volatilities, stocks' market liquidities, and funding liquidity may interact in a non-trivial fashion. To shed light on their nature and mutual influence, we model their dynamics through an MGARCH-VAR process. The model

*We thank the conference participants at the XXIII Workshop on Quantitative Finance (QFW2022), University of Rome "Tor Vergata", Rome, 31st March-1st April 2022 and the participants at the XXIV Workshop on Quantitative Finance (QFW2023), University of Cassino and Southern Lazio, Gaeta, 20th–22nd April 2023. The authors acknowledge support from the project HiDEA (Advanced Econometric Methods for High-frequency Data) financed by the Italian Ministry of Education, University and Research under the program PRIN 2017, Prot. 2017RSMPZZ and from the project PRICE (A New Paradigm for High-Frequency Finance) financed by the Italian Ministry of University and Research under the program PRIN 2022, Prot. 2022C799SX. All errors are our own.

†Alma Mater Studiorum – Università di Bologna, Dipartimento di Matematica, Piazza di Porta S. Donato, 5, 40126 Bologna, Italy.

‡Università degli Studi di Roma "Tor Vergata", Dipartimento di Economia e Finanza, Via Columbia 2, 00173 Roma, Italy. E-mail: davide.pirino@gmail.com.

is flexible enough to capture some of the well-known empirical features of financial data, such as fat-tailed distributions and conditional heteroskedasticity. Following an econometric methodology of standard practice in the realized volatility literature, the model is fitted on estimates (obtained from intra-day data through the structural model estimation) of the daily proxies for stocks' volatilities, stocks' market liquidities, and funding liquidity. On a dataset of NYSE stocks, we find significant evidence in favor of four stylized facts: (i) stocks' volatilities, stocks' market liquidities, and funding liquidity co-move; (ii) co-movements are stronger when funding liquidity dries up; (iii) stocks with lower volatility are characterized by higher market liquidity, and (iv) funding liquidity restrictions have a stronger impact on stocks' market illiquidities of high-volatility stocks.

Keywords: funding illiquidity, market illiquidity, structural estimation, market microstructure.

1 Introduction

Recent theoretical works suggest that assets' market liquidity (in its most general meaning, defined as the ease with which assets are traded) and traders' funding liquidity (the ease with which traders raise funds) are mutually reinforcing (see, e.g., Gromb and Vayanos, 2002, 2010; Geanakoplos, 2003; Brunnermeier and Pedersen, 2009, and Gârleanu and Pedersen, 2011). Low funding liquidity makes traders reluctant to take on positions, lowering stocks' market liquidity. On the other hand, low stocks' market liquidities make financiers unwilling to lend capital, reducing funding liquidity and leading to illiquidity spirals. Similar illiquidity spirals exacerbated the liquidity and credit crunch of 2007-2008 (Brunnermeier, 2009).

However, stocks' market liquidities and funding liquidity are not directly observable. Accordingly, in empirical studies, they are typically replaced by a variety of proxies, most of them based on *ad hoc* daily data (see Le and Gregoriou, 2020, for a survey on market liquidity proxies). Building upon the structural approach of Bandi et al. (2017) for stale prices (and its multivariate extension introduced in Bandi et al., 2024), this paper proposes a micro-founded structural model of price formation designed for the estimation of stocks' market liquidities and market funding liquidity. The model can be estimated by maximizing a (known analytically) likelihood function using intraday transaction prices only. The proposed framework is thus designed to deliver historical time series of stocks' market liquidities and funding liquidity (and robust estimation errors) daily. These estimated time series are revealing of new economic insights about market dynamics.

Following Glosten and Milgrom (1985), we consider a discrete-time asymmetric information framework populated by two types of traders: noise and (partially) informed agents. The arrival of an informed agent is assumed to have a constant probability π. Consequently, a noise trader has a probability of $1 - \pi$ to trade at each point in time. In addition, we postulate the existence of a (latent) price process defined as a trade signal partially known only to non-noise traders. Accordingly, these agents base their trading decisions on this process's conditional (upon their private information set) expected value. Following successful contributions of the literature (such as Hasbrouck, 2009; Brunnermeier and Pedersen, 2009), we interpret the mean absolute deviation of the transaction price from the latent price as a proxy of the stock market (ill-)liquidity. We also assume that, at each transaction, the informed trader has to pay, on top of the immediacy cost (half of the bid-ask spread) faced by all types (informed and noise) of traders, an additional cost that, following Brunnermeier and Pedersen (2009), can be interpreted as a shadow cost of capital. This cost is a natural proxy for market funding (ill-)liquidity.

With respect to the traditional asymmetric information framework (Kyle, 1985; Glosten and Milgrom, 1985; Easley and O'Hara, 1987; see Madhavan, 2000, Biais et al., 2005, Bandi et al., 2017 and Bandi et al., 2024 for a survey) our modeling approach differentiate along two major directions.

First, we replace the assumption of perfect information, according to which the informed agents have an exact knowledge of the latent price, with an assumption of partial information for the informed agents. Specifically, we assume that the informed agents only know the expected value of the latent price, given that a trend in the latent price dynamics is underway. If one of such trends occurs, the informed agents trade in the direction of the trend. Otherwise, they decide not to trade, and, as a consequence, the transaction price is not updated. On the contrary, noise traders base their decision to buy or sell on a coin toss. Such a partial information framework is warranted in contexts of high-frequency automated trading (Huang et al., 2019), which are becoming increasingly impactful[1]. This parallelism assumes that automated trading is equivalent to an informed agent that trades simultaneously on multiple assets (and in real-time) against many short

[1]Farouh and Garcia (2021) notice that "According to a recent article in The Economist, funds run by computers that follow rules set by humans account for 35% of American stock market, 60% of institutional equity assets and 60% of trading activity. According to Deutsche Bank, 90% of equity-futures trades and 80% of cash-equity trades are executed by algorithms without any human input."

trends.

The second novelty concerns the efficiency of the market. Typically, the informed agents are assumed to trade only if trading is profitable compared with the execution costs, thereby getting strictly positive per-trade net profits (See, e.g., Glosten and Milgrom, 1985, and Bandi et al., 2024). We replace this assumption by postulating that partially informed agents trade only if trading is expected to cover the execution costs, which is in accordance with a market efficiency perspective. Strictly negative per-trade losses, even for partially informed traders, are thus allowed in the model.

The assumption of partial information carries the notable advantage that the model can be written in a switching regime representation with a closed-form log-likelihood (Hamilton, 1994), whose maximization delivers daily estimates of the model parameters (and their corresponding estimation errors) exclusively based on intra-day transaction prices (sampled at sufficiently high frequency). The proposed framework provides parametric estimators of market microstructure characteristics (such as funding and stocks' market liquidity, bid-ask spreads, learning speed of the market maker, price volatility, and many others) that inherit all the nice well-known properties of the maximum likelihood estimators.

Theory indicates that stocks' volatilities, stocks' market liquidities, and funding liquidity may interact in a pretty non-trivial way. To shed light on their nature and mutual influence, in the second part of the paper, we define daily proxies for stocks' volatilities, market liquidities, and funding liquidity. We also model their dynamics through an MGARCH-VAR process. The model is flexible enough to capture some of the well-known empirical features of financial data, such as fat-tailed distributions and conditional heteroskedasticity. Following an econometric methodology of standard practice in the realized volatility literature, the model is fitted on estimates (obtained from intra-day data through the structural model) of the daily proxies for stocks' volatilities, stocks' market liquidities, and funding liquidity. For this purpose, we use a dataset of intra-day transaction prices of 150 NYSE-listed stocks. This empirical exercise reveals four stylized facts: (i) stocks' volatilities, stocks' market liquidities, and funding liquidity co-move; (ii) co-movements are stronger when funding liquidity dries up; (iii) stocks with lower volatility are characterized by higher market liquidity and (iv) stocks with higher volatility are more sensitive, in terms of their market illiquidity, to adverse shocks to funding liquidity.

The paper is organized as follows. Section 2 positions the paper in the reference literature. Section 3 introduces the micro-founded structural model. In Section 4, we define the daily proxies for funding liquidity, stocks' market liquidities and stocks' fundamental volatilities, and the related MGARCH-VAR model. Section 5 discusses the daily proxies and the MGARCH-VAR model's estimation. Section 6 illustrates some preliminary empirical evidence. In Section 7, we validate and test four conjectures concerning the dynamics of interaction between stocks' fundamental volatilities, stocks' market liquidities, and funding liquidity. Section 8 concludes the paper. The technical material is collected in the Appendices.

2 Related literature

Stocks' market liquidity is multifaceted, encompassing multiple dimensions such as traded volumes, trading intensity, and price impact (Le and Gregoriou, 2020). For instance, the effective bid-ask spread of Roll (1984) and the *Zero* estimator by Lesmond et al. (1999) measure stocks' market liquidities in terms of transaction costs, while the Amihud (2002) return-to-volume ratio and the Florackis et al. (2011) price-impact ratio, in terms of the price impact due to a change in the traded volume. Liu (2006) proposes a multi-dimensional liquidity proxy that considers volume, transaction costs, and trading speed. Goyenko et al. (2009) find a strong correlation between the *Zero* measure by Lesmond et al. (1999), the Amihud (2002) illiquidity ratio and some intraday trade-and-quote benchmarks (see also Fong et al., 2014).

Funding liquidity is less elusive, being directly related to the cost of funding. Brunnermeier (2009) and Frazzini and Pedersen (2014) measure funding liquidity via the TED spread, defined as the difference between the three-month Treasury bill rate and the three-month LIBOR (in US dollars). An increase in the TED spread signals that lenders believe default risk is increasing and funding conditions are getting tight. Gârleanu and Pedersen (2011) use the LIBOR general collateral repo interest-rate spread as a proxy for funding liquidity, while Park (2015) use the Libor-Overnight Index Swaps spread. Comerton-Forde et al. (2010) use inventory positions of NYSE specialists as funding liquidity proxy. Boudt et al. (2017) proxy for the aggregate daily funding liquidity of S&P 500 stocks using a volume-weighted average of their stock loan rates. For the Chinese stock market, Qian et al. (2014) use the monthly change in the number of market participants. The authors argue that new market participants bring new funds, thereby increasing funding liquidity. Other low-frequency funding liquidity measures can be found in Fontaine and Garcia (2012), Hu et al. (2013), and Golez et al. (2017).

Regarding the relationships between market liquidity and funding liquidity, Hameed et al. (2010) show that changes in the value of collateralized equities affect market liquidity. Comerton-Forde et al. (2010) show that liquidity-supplier financing constraints affect time-variation in market liquidity. Mancini-Griffoli and Ranaldo (2011) compare the effect of secured versus unsecured borrowing by arbitrageurs and conclude that funding liquidity affects market liquidity. Mancini et al. (2013) analyze liquidity in foreign exchange markets and provide evidence of liquidity spirals between traders' funding liquidity and market-wide foreign exchange liquidity. Boudt et al. (2017) show that the effect of market liquidity on funding liquidity differs in stabilizing and destabilizing regimes. Farouh and Garcia (2021) study the effect of funding liquidity conditions on time variation in bid-ask spreads in the last 30 years. They find that three-quarters of the firms are significantly impacted by funding liquidity. Macchiavelli and Zhou (2022) focus on the dealer's perspective and find that the dealer's repo trading terms affect the dealer's liquidity provision in securities markets.

3 Micro-founded structural model

In what follows, we interchangeably indicate with $\{x_t \mid t = 0, 1, 2, \dots\}$, $\{x_t\}$ or, simply, x a discrete-time stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathsf{P})$. Our partial information framework features four (main) stochastic processes (in discrete time): a latent log-price, $\{e_t\}$, the market maker's mid-quote, indicated with $\{m_t\}$, the transaction log-price, $\{p_t\}$ and a trade signal, $\{g_t\}$. We consider three kinds of agents: noise agents, partially informed agents (sometimes referred to simply as *informed traders/agents*), and the market maker. The latent price process $\{e_t\}$ is unknown to all agents. The mid-quote and the transaction price processes, respectively $\{m_t\}$ and $\{p_t\}$, are public information. Finally, the trade signal $\{g_t\}$ is known only by informed agents. We first introduce a semiparametric model that, with some suitable parametric assumptions, allows for the definition of a maximum likelihood (ML) estimator.

3.1 The semiparametric model

We assume that the trade signal, whose role is clarified later in this section, follows a Markov process with values in $\{-g, 0, g\}$ for some real $g > 0$. We denote with $\mathcal{G}_t \doteq \sigma\left\{g_r, m_r \,\middle|\, 0 \le r \le t\right\}$ the filtration generated, at time t, by the processes $\{g_t\}$ and $\{m_t\}$. In the model's logic, this filtration is interpreted as the private information set of the partially informed agents. We postulate that, for each t, $g_t = \bar{e}_t - m_t$, where

$$\bar{e}_t \doteq \mathsf{E}\{e_t \mid m_t, \mathcal{G}_{t-1}\}, \quad t = 1, 2, \ldots, T, \tag{1}$$

denotes the latent price expected by the partially informed agents at time t, and T denotes the time horizon. Accordingly, the trade signal coincides with the signed profit, gross of the execution costs, expected by partially informed agents at time t. Noticing that $\bar{e}_t = g_t + m_t$, the information set of the partially informed agents can equivalently be defined as $\bar{\mathcal{E}}_t \doteq \sigma\{\bar{e}_r, m_r \mid 0 \leq r \leq t\}$. The terminology *partially informed* is warranted to stress the fact that this kind of players has access to only the expected latent price $\{\bar{e}_t\}$, but not the latent price process $\{e_t\}$.

We assume that $\{\bar{e}_t\}$ evolves according to the recursive equation

$$\bar{e}_t = \bar{e}_{t-1} + \eta_t, \quad t = 1, 2, \ldots, T, \tag{2}$$

with $\bar{e}_0 = 0$ and where the shocks η's are such that

$$\mathsf{E}\left\{\eta_t \,\middle|\, \bar{\mathcal{E}}_{t-1}\right\} = 0, \qquad \mathsf{E}\left\{\eta_t^2\right\} = \sigma_\eta^2 \qquad \text{and} \qquad \sigma_\eta \geq 0. \tag{3}$$

The innovation process $\{\eta_t\}$ is thus a (stationary) martingale difference with respect to $\bar{\mathcal{E}}_{t-1}$. Accordingly, $\{\bar{e}_t\}$ follows a random walk.

Even though the market maker does not know $\{\bar{e}_t\}$, in line with previous studies (see, for example, Hasbrouck and Ho, 1987; Amihud and Mendelson, 1987; Bandi et al., 2017), we assume that it can partially infer, from the order flow, its value. Borrowing from Bandi et al. (2017), we implement this model feature by assuming that $\{m_t\}$ evolves as

$$m_t = m_{t-1} + \delta(\bar{e}_t - m_{t-1}) + (1 - \delta)\zeta_t, \quad t = 1, 2, \ldots, T, \tag{4}$$

where

$$\mathsf{E}\left\{\zeta_t \,\middle|\, \bar{\mathcal{E}}_{t-1}\right\} = 0, \qquad \mathsf{E}\left\{\zeta_t^2\right\} = \sigma_\zeta^2, \qquad \text{and} \qquad \sigma_\zeta \geq 0. \tag{5}$$

Moreover, we assume that the innovations $\{\eta_t\}$ and $\{\zeta_t\}$ are orthogonal, that is

$$\mathsf{E}\{\eta_t \zeta_t\} = 0, \quad t = 1, 2, \ldots, T. \tag{6}$$

The parameter δ in equation (4) can be interpreted as the learning speed of the market maker: the closer the δ to one (resp. to zero), the faster (resp. the slower) the market maker adjusts m_t toward $\bar{e}_t$.

In the spirit of Glosten and Milgrom (1985), we assume that at time $t > 0$, either a noise agent or an informed agent arrives. The probability of arrival of an informed agent is denoted with π. A noise agent either sells at $p_t = m_t - s$ or buys at $p_t = m_t + s$ with equal probability, where $s > 0$ denotes the half of the bid-ask spread. An informed agent sells at $p_t = m_t - s$ if $g_t = -g$, buys at $p_t = m_t + s$ if $g_t = g$ and refrains from trading in the market whenever $g_t = 0$. In the latter case, trading does not occur, and the transaction price does not update, i.e. $p_t = p_{t-1}$ (borrowing from the nomenclature of Bandi et al., 2020, we will refer to this event as a *stale price*). As initial value we set $p_0 = 0$.

Let us denote by $\{I_t\}$ and $\{B_t\}$ two Bernoulli processes with success probability π and 1/2, respectively. The two processes are assumed to be independent of each other and of $\{\eta_t\}$ and $\{\zeta_t\}$. They are assumed to control the arrival of noise and informed traders and their trading decisions. More specifically, at a generic time t, the event $I_t = 0$ corresponds to the arrival of a noise trader while $I_t = 1$ to that of an informed agent. The decision at time t, to sell or buy is based, for a noise trader, upon the value of B_t as illustrated by the following scheme:

$$I_t = 0 \rightarrow \begin{cases} B_t = 0 \rightarrow p_t = m_t - s & \text{(sale)}, \\ B_t = 1 \rightarrow p_t = m_t + s & \text{(buy)}. \end{cases} \tag{7}$$

For an informed agent, the trading decisions depend on the trade signal, g_t, according to the following scheme:

$$I_t = 1 \rightarrow \begin{cases} g_t = -g \rightarrow p_t = m_t - s & \text{(sale)}, \\ g_t = 0 \rightarrow \quad p_t = p_{t-1} & \text{(no trade)}, \\ g_t = g \rightarrow \quad p_t = m_t + s & \text{(buy)}. \end{cases} \tag{8}$$

Equations (7–8) provide the data generating process (DGP) of the transaction (log) price process $\{p_t\}$. Recalling that the informed agents trade exclusively whenever $g_t \neq 0$, we interpret g as the expected *gross* profit for an informed trade. The corresponding expected *net* profit is assumed to be, at any given time t, equal to $g - c$, where c denotes the per-trade execution cost for an informed agent. We set $c \doteq f + s$, where $f > 0$ denotes the per-trade cost of funding borne by the informed agents (the "shadow cost of capital", Brunnermeier and Pedersen, 2009). Then, we identify f by imposing a market efficiency assumption. We assume that $c = g$, which amounts to postulating a zero net expected profit for informed agents[2]. Accordingly, we get that $f = g - s$, whence we impose $s < g$ for f to be meaningful.

To complete our semiparametric framework, we assume that the latent price $\{e_t\}$ evolves as $e_t = e_{t-1} + \eta_t + \varepsilon_t$, where $e_0 = 0$ and $\{\varepsilon_t\}$ is an innovation process, independent of

[2]This assumption does not contradict the possibility for an informed trader to gain using her private information. Perhaps it imposes a global market efficiency across the collection of all the trades.

$\{(\eta_t, \zeta_t, I_t, B_t)\}$ and such that $\mathsf{E}\{\varepsilon_t\} = 0$, for $t = 1, 2, \ldots, T$. For $\{e_t\}$ to be consistent with the definition of the triplet $\{(\bar{e}_t, m_t, p_t)\}$, it suffices to verify that $\mathsf{E}\left\{e_t \mid m_t, \bar{\mathcal{E}}_{t-1}\right\} = \bar{e}_t$, as required by (1). This is immediate, since

$$e_t = \sum_{j=1}^{t} \eta_j + \sum_{j=1}^{t} \varepsilon_j = \bar{e}_t + \bar{\varepsilon}_t, \tag{9}$$

where $\bar{\varepsilon}_t \doteq \sum_{j=1}^{t} \varepsilon_j$, for $t \geq 0$. Whence

$$\mathsf{E}\left\{e_t \mid m_t, \bar{\mathcal{E}}_{t-1}\right\} = \mathsf{E}\left\{\bar{e}_t \mid m_t, \bar{\mathcal{E}}_{t-1}\right\} + \mathsf{E}\left\{\bar{\varepsilon}_t \mid m_t, \bar{\mathcal{E}}_{t-1}\right\} = \bar{e}_t + \mathsf{E}\{\bar{\varepsilon}_t\} = \bar{e}_t + 0 = \bar{e}_t,$$

where we replaced $\mathsf{E}\left\{\bar{\varepsilon}_t \mid m_t, \bar{\mathcal{E}}_{t-1}\right\}$ with $\mathsf{E}\{\bar{\varepsilon}_t\}$ because $\{\varepsilon_t\}$ is independent of $\{(\eta_t, \zeta_t)\}$.

3.2 Parametric specification and closed-form log-likelihood

Assuming that the econometrician observes only $\{p_t\}$, it is possible to derive an ML estimator of the model parameters as soon as a parametric model for the vector of innovations $\{(\eta_t, \zeta_t)\}$ and for the trade signal $\{g_t\}$ is specified. In doing this, the parametric choice ought to be consistent with the fact that 1) $\{(\eta_t, \zeta_t)\}$ is a stationary orthogonal martingale difference and 2) $\{g_t\}$ is a Markov process with values in $\{-g, 0, g\}$, for some $g > 0$. The following theorem describes how to obtain what we are looking for.

Theorem 3.1. *Let the following assumptions hold.*

$\mathcal{A}_0$ - *The innovations $\{\eta_t\}$ and $\{\zeta_t\}$ are such that*

$$\eta_t \doteq \nu_t + \gamma_\eta u_t \qquad \text{and} \qquad \zeta_t \doteq \nu_t - \gamma_\zeta u_t, \quad t = 1, \ldots, T, \tag{10}$$

where

$$u_t \doteq g_t - \omega g_{t-1}, \qquad \gamma_\eta \doteq \frac{a}{\omega(a+b)}, \qquad \gamma_\zeta \doteq \frac{b}{\omega(a+b)}, \tag{11}$$

with $a > 0$ and $b > 0$ real positive parameters and $\omega \doteq 1 - \delta$.

$\mathcal{A}_1$ - *The process $\{g_t\}$ is a Markov process taking values in $\{-g, 0, g\}$, with $g > 0$ a real constant given by*

$$g \doteq \sqrt{a+b} \cdot \frac{\omega}{\sqrt{1-\omega^2}} \cdot \sqrt{\frac{2-\omega-\psi}{1-\psi}}, \tag{12}$$

where $\psi \in (0,1)$. The transition matrix T of $\{g_t\}$, whose generic entry is

$$\mathsf{T}_{ij} \doteq \mathsf{P}\left\{g_t = \ell_i g \mid g_{t-1} = \ell_j g\right\},$$

for $i, j = 1, 2, 3$, where $(\ell_1, \ell_2, \ell_3) \doteq (-1, 0, 1)$ and $t = 1, \dots, T$, is given by

$$\mathsf{T} \doteq \begin{pmatrix} \omega & 1-\omega & 0 \\ (1-\psi)/2 & \psi & (1-\psi)/2 \\ 0 & 1-\omega & \omega \end{pmatrix}. \tag{13}$$

$\mathcal{A}_2$ *- The distribution of the initial value g_0 is*

$$\mathsf{P}\{g_0 = -g\} \doteq \mathsf{P}\{g_0 = g\} \doteq (1-\psi)/(2(2-\omega-\psi)), \qquad \mathsf{P}\{g_0 = 0\} = 1-2\mathsf{P}\{g_0 = g\}. \tag{14}$$

$\mathcal{A}_3$ *- The process $\{v_t\}$ is a Gaussian white noise with variance*

$$\mathsf{E}\left\{v_t^2\right\} \doteq \sigma_v^2 \doteq (ab)/(a+b), \quad t = 1, \dots, T. \tag{15}$$

Under Assumptions from $\mathcal{A}_0$ to $\mathcal{A}_3$, it follows that equations (3) and (5-6) hold with $\mathsf{E}\left\{\eta_t^2\right\} = a$ and $\mathsf{E}\left\{\zeta_t^2\right\} = b$. In particular, it holds that $\bar{e}_t - m_t = g_t$, for all t.

Proof. See Appendix A.

The model for $\{p_t\}$ given in (2), (4), and (7–8), with $\{(\eta_t, \zeta_t)\}$ defined as in Theorem 3.1, turns out to be parameterized by the vector $\tilde{\theta} \doteq (s, a, b, \delta, \pi, \psi)^\top$, where ψ represents a probability. While the parameters $\{a, b, \delta, \pi, \psi\}$ are variation-free, the parameter s is subject to $s < g$, where g is the function of $\{a, b, \delta, \psi\}$ defined in (12). Noticing that $(\sigma_\eta^2, \sigma_\zeta^2) = (a, b)$, and defining $\phi \doteq f/s$, we can reparameterize $\{p_t\}$, more conveniently, in terms of the vector

$$\theta \doteq (\phi, \sigma_\eta, \sigma_\zeta, \delta, \pi, \psi)^\top.$$

Unlike the parameters in $\tilde{\theta}$, the parameters in θ are all variation-free[3].

The parameter ϕ can be interpreted as a normalized funding cost and plays a crucial role in the definition of our funding ill-liquidity proxy[4]. Recalling the formula of the transition matrix T of the trade signal $\{g_t\}$ given in (13), the parameter ψ can be interpreted

[3]To recover $\theta = (\phi, \sigma_\eta, \sigma_\zeta, \pi, \delta, \psi)^\top$ from $\tilde{\theta} = (s, a, b, \delta, \pi, \psi)^\top$, we first replace (a, b) with $(\sigma_\eta^2, \sigma_\zeta^2)$ as an application of the last part of Theorem 3.1. Then, applying definition (12), and using $\omega = 1 - \delta$, we compute g as

$$g = \sqrt{\sigma_\eta^2 + \sigma_\zeta^2} \cdot \frac{1-\delta}{\sqrt{\delta\,(2-\delta)}} \cdot \sqrt{\frac{1+\delta-\psi}{1-\psi}}.$$

Finally, recalling that $g = f + s$, we compute f as $f = g - s$, and we set $\phi = f/s$. To recover $\tilde{\theta}$ from θ, we first replace $(\sigma_\eta^2, \sigma_\zeta^2)$ with (a, b). Then, we compute g applying (12), with $\omega = 1 - \delta$. Finally, applying $\phi = f/s$, and $g = f + s$, we derive s as $s = g/(1 + \phi)$.

[4]See Section 4.

as the persistence of unprofitable (expected) trades, in formula $\psi = \mathsf{P}\{g_t = 0 \mid g_{t-1} = 0\}$. On the other hand, the parameter ω in T measures the persistence of the profitable trades, that is $\omega = \mathsf{P}\{g_t = -g \mid g_{t-1} = -g\} = \mathsf{P}\{g_t = g \mid g_{t-1} = g\}$. Since $\omega = 1 - \delta$, the persistence of the profitable trades is a decreasing function of the learning speed of the market maker. We now provide the closed form of the distribution of $\{p_t\}$, given the information set of the econometrician. This is assumed to be the filtration generated by transaction prices, that is

$$\mathcal{F}_t \doteq \sigma\{p_t, p_{t-1}, \ldots, p_0\}, \quad t = 1, \ldots, T, \tag{16}$$

having further assumed that $p_0 = 0$. First, we need to introduce some notations.

Let $\{s_t\}$, with $s_t \in \{-s, 0, s\}$, denote the sequence of immediacy costs, with the convention that $s_t = 0$ if no trades occur at time t. Let t_q denote the q-th point in time in which a price change occurs, with the convention that $t_0 = 0$. For $q = 0, 1, \ldots$ and for $t_q < t \le t_{q+1}$, let $P_{t-1}(s_t, g_t, s_{t_q}, g_{t_q})$ denote the probability mass function of $(s_t, g_t, s_{t_q}, g_{t_q})$ conditional on $\mathcal{F}_{t-1}$, with the convention that $(s_{t_0}, g_{t_0}) = (s_0, g_0) \doteq (0, 0)$. Let $S \doteq \{-s, 0, s\}$ and $G \doteq \{-g, 0, g\}$. Let $\mathbf{1}(\,\cdot\,)$ denote the indicator function. Let $f_G(x; \mu, \lambda)$ denote the Gaussian density with mean μ and variance λ. Finally, let $f_{t-1}(p_t; \theta)$ denote the probability density function of p_t given $\mathcal{F}_{t-1}$. The following theorem provides the closed form of $f_{t-1}(p_t; \theta)$.

Theorem 3.2. *In the model for $\{p_t\}$ given in (2), (4), and (7–8), with $\{(\eta_t, \zeta_t, g_t)\}$ defined as in Theorem 3.1, it holds, for $q = 0, 1 \ldots$ and for $t_q < t \le t_{q+1}$, that*

$$\begin{aligned} f_{t-1}(p_t; \theta) \;=\; & \sum_{s_t = 0,\, (g_t, s_{t_q}, g_{t_q}) \in G\times S\times G} P_{t-1}(s_t, g_t, s_{t_q}, g_{t_q}) \cdot \mathbf{1}(p_t = p_{t-1}) + \\ & \sum_{s_t \neq 0,\, (g_t, s_{t_q}, g_{t_q}) \in G\times S\times G} P_{t-1}(s_t, g_t, s_{t_q}, g_{t_q}) \cdot \mathbf{1}(p_t \neq p_{t-1}) \cdot f_G(p_t; \mu_t, \lambda_t), \end{aligned} \tag{17}$$

with $\mu_t \doteq p_{t_q} + s_t + \rho g_t - s_{t_q} + \xi g_{t_q}$, $\rho \doteq \delta/\omega - \gamma_\zeta$, $\xi \doteq \gamma_\zeta \omega$, and $\lambda_t \doteq (t - t_q)\sigma_\nu^2$.

Proof. See Appendix A.

The expression of $f_{t-1}(p_t; \theta)$ given in equation (17) delivers a discrete/continuous mixture distribution, where the discrete components are the (degenerate) components of the form $\mathbf{1}(p_t = p_{t-1})$, and the continuous components are the Gaussian components of the form $\mathbf{1}(p_t \neq p_{t-1}) \cdot f_G(p_t; \mu_t, \lambda_t)$. The discrete components model the stale prices, $p_t = p_{t-1}$, while the continuous components model the price changes. The mixture weights, which are given by $P_{t-1}(s_t, g_t, s_{t_q}, g_{t_q})$, for $(s_t, g_t, s_{t_q}, g_{t_q}) \in S\times G\times S\times G$, can be computed via Bayesian forward recursion, as described in Appendix B.[5] Since $f_{t-1}(p_t; \theta)$ has a closed-form expression,

[5]The model of $\{p_t\}$ turns out to be a switching regime model (see Hamilton, 1994, Chapter 22).

the prediction-error decomposition $\sum_{t=1}^{T} \log f_{t-1}(p_t;\theta)$ of the log-likelihood of θ admits a closed form too.

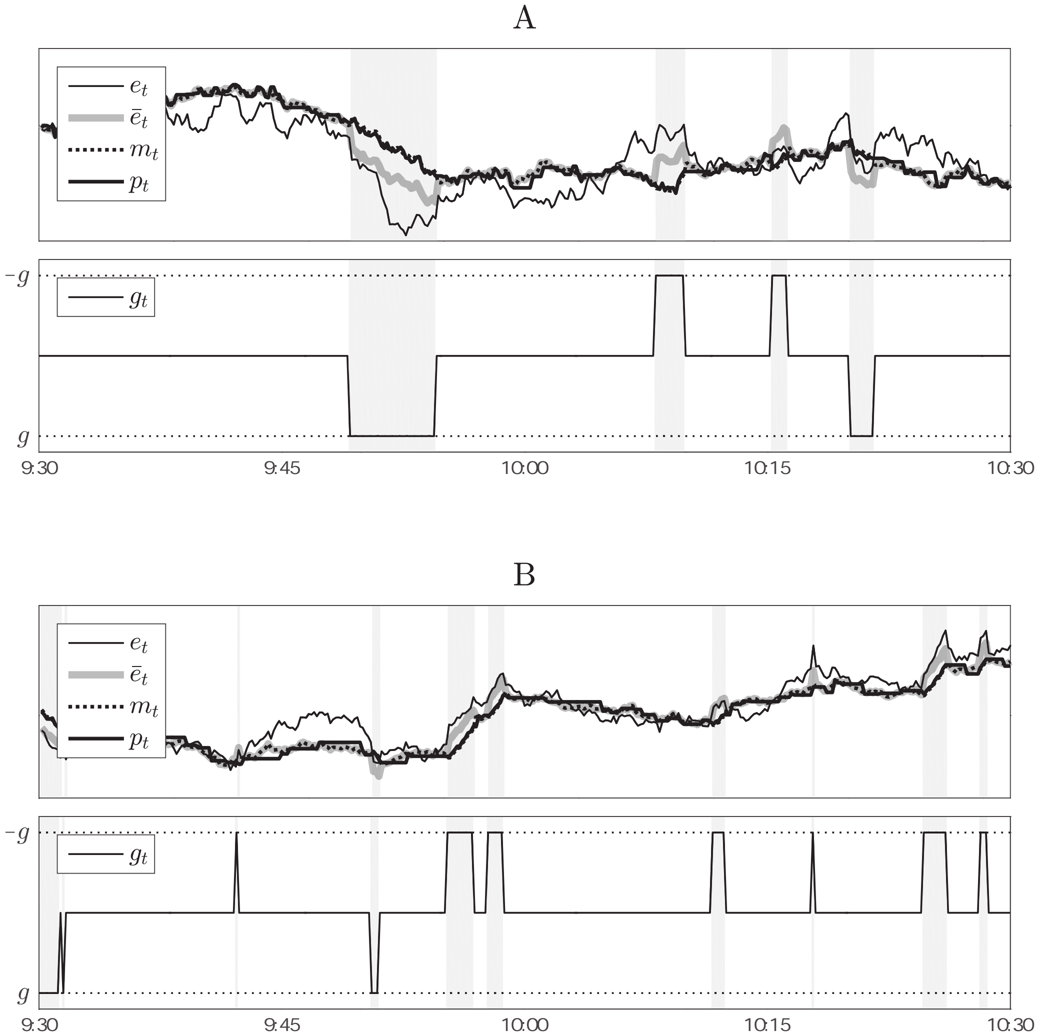

Figure 1: Simulated trajectories of the micro-founded structural model. In Panel A the parameters used are $(\phi, \sigma_\eta, \sigma_\zeta, \delta, \pi, \psi) = (9, 0.03, 0.02, 0.10, 0.5, 0.98)$, while, in Panel B, we set $(\phi, \sigma_\eta, \sigma_\zeta, \delta, \pi, \psi) = (9, 0.03, 0.02, 0.20, 0.8, 0.98)$. In both panels the latent price $\{e_t\}$ is generated applying equation (9), with $\{\bar{\varepsilon}_t\}$ modeled as an AR(1) process.

To better illustrate the dynamics of the model, we report in Figure 1 two simulated trajectories (see Panel A and Panel B, respectively). The grey vertical bands highlight the profitable trends. The parameter values of the two DGPSs are the same, except for the couple (δ, π) which is $(\delta, \pi) = (0.10, 0.50)$ in Panel A and $(\delta, \pi) = (0.20, 0.80)$ in Panel B. The latent price $\{e_t\}$ is obtained by generating the process $\{\bar{\varepsilon}_t\}$ in equation (9) from a zero-mean

AR(1) process.

The persistence of the profitable trends, measured by $\omega = 1 - \delta$, is more significant in Panel A, which results in wider grey bands. The probability of arrival of an informed agent, π, is more significant in Panel B, as witnessed by the longer periods in which the price stays constant. Within the grey bands, staleness disappears (the price moves at each point) as the informed agents (the only ones responsible for price staleness) keep active during profitable trends. Recalling that informed agents trade exclusively in the direction of the trends, we have that $\{p_t\}$ follows $\{\bar{e}_t\}$ within the grey bands.

4 Daily proxies of volatility and illiquidity

We work under the assumption that funding liquidity has a systematic nature or, in other words, that its across-stock variation is negligible. As a consequence, a consistent estimation of this market feature requires a multivariate version of the model, something that we obtain by simply replicating N times the DGP described in Sections 3 and 3.2, where N is the number of stocks available for estimation. Let $d = 1, \ldots, D$ be a discrete index that runs across the day in the sample. The model parameters for the DGP of the i-th process, with $i = 1, \ldots, N$, shall be denoted as

$$\theta_d^{(i)} \doteq (\phi_d^{(i)}, \sigma_{\eta,d}^{(i)}, \sigma_{\zeta,d}^{(i)}, \pi_d^{(i)}, \delta_d^{(i)}, \psi_d^{(i)})^\top. \tag{18}$$

Accordingly, the corresponding transaction log-price time series is denoted as $\left\{p_{d,t}^{(i)}\right\}$, with the convention that $p_{d,t}^{(i)}$ denotes the transaction log-price of the i-th stock at the t-th time instant of the day d. An identical notation shall be used for all the other model time series.

Within the structural framework, a natural proxy for stock volatility is the daily volatility of the latent log price expected by the informed agents. Following this idea, we define

$$\varsigma_d^{(i)} \doteq \varsigma_{\eta,d}^{(i)} \doteq \sqrt{T}\sigma_{\eta,d}^{(i)},$$

as the i-th stock fundamental volatility for the d-th day.

To derive a proxy for market illiquidity, in line with the theoretical framework of Brunnermeier and Pedersen (2009), we proxy, at an intraday time-scale, stock market illiquidity as

$$\Lambda_{d,t}^{(i)} \doteq \left|p_{d,t}^{(i)} - \bar{e}_{d,t}^{(i)}\right|,$$

i.e. the absolute deviation of the transaction log-price, $p_{d,t}^{(i)}$, from the latent log-price expected by the informed agents, $\bar{e}_{d,t}^{(i)}$. At a daily time scale, the proxy for the stock market illiquidity is defined as $\Lambda_d^{(i)} \doteq T^{-1} \sum_{t=1}^{T} \Lambda_{d,t}^{(i)}$, that is the intra-daily average.

Concerning the identification of a proxy for daily funding illiquidity, we remind that, in the micro-founded structural model, funding and transaction costs are represented, for a given stock i and a given day d, by the parameters $f_d^{(i)}$ and $s_d^{(i)}$, respectively. In this context, a funding cost of, say, one basis point over a half bid-ask spread of, say, one basis point is expected to be associated with a more liquid state of the world than that of a funding cost of, still, one basis point over a half bid-ask spread of, say, ten basis points. For this reason, we adopt, as a funding illiquidity proxy at a stock level, the normalized funding cost

$$\phi_d^{(i)} \doteq \frac{f_d^{(i)}}{s_d^{(i)}}.$$

On top of this, we assume a value of $\phi_d^{(i)}$ common across all stocks, that we denote with Φ_d, which means that we work under the assumption that

$$\phi_d^{(i)} = \Phi_d, \text{ for all } i = 1, \ldots, N. \tag{19}$$

The volatility and illiquidity proxies are, by construction, positive quantities. Hence, we are allowed to collect them in a vector of logarithms

$$Y_d \doteq \left(\log \varsigma_d^{(1)}, \ldots, \log \varsigma_d^{(N)}, \log \Lambda_d^{(1)}, \ldots, \log \Lambda_d^{(N)}, \log \Phi_d\right)^\top, \tag{20}$$

whose length is $M \doteq 2N + 1$. We postulate the following assumption concerning the dynamics of the vector time series defined in equation (20).

Assumption 4.1. *The vector time series $\{Y_d\}$ defined in* (20) *follows a VAR(L) process (Hamilton, 1994), that is*

$$Y_d = c_0 + C_1 Y_{d-1} + C_2 Y_{d-2} + \cdots + C_L Y_{d-L} + U_d, \tag{21}$$

where $L > 0$ is a given lag index, the vector of innovations $U_d \doteq \left(U_d^{(1)}, \ldots, U_d^{(M)}\right)^\top$ is such that $\mathsf{E}\left\{U_d \,\middle|\, \mathcal{Y}_{d-1}\right\} = 0$ (martingale difference) and where $\mathcal{Y}_d$ denotes, for all d, the σ-algebra defined as $\mathcal{Y}_d \doteq \sigma\{Y_r \mid r = -L+1, \ldots, d\}$.

To allow for conditional heteroskedasticity, we model the conditional var-covar matrix $H_d \doteq \mathsf{E}\left\{U_d U_d^\top \,\middle|\, \mathcal{Y}_{d-1}\right\}$ as a MGARCH process (Bauwens et al., 2006; Silvennoinen and Teräsvirta, 2009). In particular, we adopt the following Diagonal VECH specification (Bollerslev et al., 1988)

$$H_d = W \odot (\mathbb{I}_{M,M} - A - B) + A \odot U_{d-1} U_{d-1}^\top + B \odot H_{d-1}, \tag{22}$$

where $\odot$ denotes the element-wise matrix product, $\mathbb{I}_{M,M}$ denotes the $M \times M$-dimensional matrix whose entries are all ones, W, A and B are $M \times M$ parameter matrices. In particular, the conditional variance of the i-th element $U_d^{(i)}$ follows the GARCH(1,1) model

$$H_d^{(i,i)} = W_{(i,i)}(1 - A_{(i,i)} - B_{(i,i)}) + A_{(i,i)} \left(U_{d-1}^{(i)}\right)^2 + B_{(i,i)} H_{d-1}^{(i,i)},$$

where (i, j) denotes typical matrix element.

For $\{H_d\}$ to be positive definite we impose that the matrices A, B, and $W \odot (\mathbb{I}_{M,M} - A - B)$ are positive definite (Ding and Engle, 2001). A consequence of this is that

$$A_{(i,i)} + B_{(i,i)} < 0, \quad \text{for} \quad i = 1, \ldots, M. \tag{23}$$

Under additional regularity conditions (Bollerslev et al., 1988; Boussama et al., 2011), $\{U_d\}$ is strictly and weakly stationary with unconditional second moment

$$\mathsf{E}\left\{U_d\, U_d^\top\right\} = W.$$

The number of distinct parameters in each of the H_d's is $3M(M + 1)/2$, where $M(M + 1)/2$ corresponds to the number of distinct entries in each of the parameter matrices W, A, and B. To reduce the number of parameters, we take advantage of the partition $Y_d \doteq \left(Y_d^{(\varsigma)\top}, Y_d^{(\lambda)\top}, Y_d^{(\phi)}\right)^\top$, with subvectors defined as $Y_d^{(\varsigma)\top} \doteq \left(\log \varsigma_d^{(1)}, \ldots, \log \varsigma_d^{(N)}\right)^\top$, $Y_d^{(\lambda)\top} \doteq \left(\log \Lambda_d^{(1)}, \ldots, \log \Lambda_d^{(N)}\right)^\top$, and $Y_d^{(\phi)} \doteq \log \Phi_d$. Following Billio et al. (2006), we force the dynamic parameters to be equal across homogeneous variance-covariance processes, getting block structured A and B of the form

$$A = \begin{pmatrix} \alpha_{\varsigma,\varsigma}\mathbb{I}_{N,N} & \alpha_{\varsigma,\lambda}\mathbb{I}_{N,N} & \alpha_{\varsigma,\phi}\mathbb{I}_{N,1} \\ \alpha_{\lambda,\varsigma}\mathbb{I}_{N,N} & \alpha_{\lambda,\lambda}\mathbb{I}_{N,N} & \alpha_{\lambda,\phi}\mathbb{I}_{N,1} \\ \alpha_{\phi,\varsigma}\mathbb{I}_{1,N} & \alpha_{\phi,\lambda}\mathbb{I}_{1,N} & \alpha_{\phi,\phi}\mathbb{I}_{1,1} \end{pmatrix}, \qquad B = \begin{pmatrix} \beta_{\varsigma,\varsigma}\mathbb{I}_{N,N} & \beta_{\varsigma,\lambda}\mathbb{I}_{N,N} & \beta_{\varsigma,\phi}\mathbb{I}_{N,1} \\ \beta_{\lambda,\varsigma}\mathbb{I}_{N,N} & \beta_{\lambda,\lambda}\mathbb{I}_{N,N} & \beta_{\lambda,\phi}\mathbb{I}_{N,1} \\ \beta_{\phi,\varsigma}\mathbb{I}_{1,N} & \beta_{\phi,\lambda}\mathbb{I}_{1,N} & \beta_{\phi,\phi}\mathbb{I}_{1,1} \end{pmatrix}, \tag{24}$$

where α and β denote scalar parameters. For instance, the conditional covariance between any stock market log-illiquidity, $\log \Lambda_d^{(N+i)}$, $i = 1, \ldots, N$, and any stock fundamental log-volatility, $\log \varsigma_d^{(j)}$, $j = 1, \ldots, N$, takes the form

$$H_d^{(N+i,j)} = W_{(N+i,j)}(1 - \alpha_{\lambda,\varsigma} - \beta_{\lambda,\varsigma}) + \alpha_{\lambda,\varsigma} U_{d-1}^{(N+i)} U_{d-1}^{(j)} + \beta_{\lambda,\varsigma} H_{d-1}^{(N+i,j)},$$

where the scalar parameters $\alpha_{\lambda,\varsigma}$ and $\beta_{\lambda,\varsigma}$ do not depend on (i, j). Irrespective of N, the distinct entries of A and B are now 12 in total, as we have 6 distinct scalar entries, $(\alpha_{\varsigma,\varsigma}, \alpha_{\lambda,\varsigma}, \alpha_{\phi,\varsigma}, \alpha_{\lambda,\lambda}, \alpha_{\phi,\lambda}, \alpha_{\phi,\phi})$, in A, plus 6 distinct scalar entries, $(\beta_{\varsigma,\varsigma}, \beta_{\lambda,\varsigma}, \beta_{\phi,\varsigma}, \beta_{\lambda,\lambda}, \beta_{\phi,\lambda}, \beta_{\phi,\phi})$, in B.

Having implemented conditional heteroskedasticity, we implement fat-tailed innovations by modeling the conditional distribution of the innovations $\{U_d\}$ as a M-variate t-distribution, that is

$$f^{(U)}(U_d|\mathcal{Y}_{d-1}) \doteq \frac{\Gamma[(M+\kappa)/2]}{\Gamma[\kappa/2](\kappa\pi)^{M/2}|S_d|^{1/2}}\left(1+\frac{1}{\kappa}U_d^\top S_d^{-1}U_d\right)^{-(M+\kappa)/2}, \tag{25}$$

where $\Gamma[\,\cdot\,]$ is the gamma function, S_d is, for all d, a positive definite matrix, and $\kappa \in (0,\infty)$ is the kurtosis parameter (Kotz and Nadarajah, 2004). For $\kappa \longrightarrow \infty$, the distribution tends to the multivariate standard normal distribution. If $\kappa > 2$ the conditional second moment of each of the U_d's is finite and equal to $S_d\kappa/(\kappa-2)$, in which case $S_d = H_d(\kappa-2)/\kappa$.

5 Estimation

An inherent benefit of structural modeling lies in its ability to provide daily realized measures of market and funding liquidity (as measured by their proxies defined in Section 4) based on high-frequency intra-day transaction data. These estimates are, in nature, parametric, so they differ in this aspect from the typical approach used in the realized volatility literature (where, very frequently, integrated volatility is estimated via functionals of the intra-day time series of log-returns). We illustrate, in what follows, how such estimates can be obtained. The resulting estimated daily realized time series of volatility and (market plus funding) illiquidity constitute the inputs for the dynamical model postulated in Assumption 4.1, which, once fitted on market data, we deem to be revealing on the self and mutual influence among these market characteristics. We also comment on the estimation of the MGARCH-VAR in this section.

To begin, let $\hat{\theta}_d^{(i)} \doteq (\hat{\phi}_d^{(i)}, \hat{\sigma}_{\eta,d}^{(i)}, \hat{\sigma}_{\zeta,d}^{(i)}, \hat{\pi}_d^{(i)}, \hat{\delta}_d^{(i)}, \hat{\psi}_d^{(i)})^\top$ denote, for $i = 1,\ldots,N$ and $d = 1,\ldots,D$, the univariate ML estimator of the vector of parameters $\theta_d^{(i)}$, defined in equation (18).

For any given $i = 1,\ldots,N$, the numerical computation of the vector time series $\left\{\hat{\theta}_d^{(i)} \,\middle|\, d = 1,\ldots,D\right\}$ is straightforward, as the prediction-error decomposition of the log-likelihood of each of the $\theta_d^{(i)}$'s has the closed form derived from equation (17). After computing the $\hat{\theta}_d^{(i)}$'s we estimate the $\varsigma_d^{(i)}$'s through $\hat{\varsigma}_d^{(i)} \doteq \sqrt{T}\hat{\sigma}_{\eta,d}^{(i)}$.

The daily stock market illiquidity proxy, $\Lambda_d^{(i)} = T^{-1}\sum_{t=1}^T \Lambda_{d,t}^{(i)} = T^{-1}\sum_{t=1}^T \left|p_{d,t}^{(i)} - \bar{e}_{d,t}^{(i)}\right|$, depends on the latent process $\left\{\bar{e}_{d,t}^{(i)}\right\}$, which is, by definition, not available to the researcher. Nevertheless, its intraday fixed interval smoothing mean, namely $\mathsf{E}\left\{\bar{e}_{d,t}^{(i)} \,\middle|\, \mathcal{F}_T\right\}$ (where $\mathcal{F}_T$, defined in equation (16), is the information set available to the researcher), computed at the ML estimator $\hat{\theta}_d^{(i)}$, can be estimated. We prove this in Appendix C, where we define an estimator $\hat{e}^{(i)}_{d,t}$ of $\mathsf{E}\left\{\bar{e}^{(i)}_{d,t} \mid \mathcal{F}_T\right\}$. Having at our disposal the time series $\left\{\hat{e}^{(i)}_{d,t}\right\}$, we can estimate $\left\{\Lambda^{(i)}_d\right\}$ through

$$\hat{\Lambda}^{(i)}_d \doteq T^{-1} \sum_{t=1}^{T} \hat{\Lambda}^{(i)}_{d,t} \doteq T^{-1} \sum_{t=1}^{T} \left|p^{(i)}_{d,t} - \hat{e}^{(i)}_{d,t}\right|,$$

Finally, to estimate the Φ_d's we simply compute, for each day $d = 1, \ldots, D$, the across-stock median of the estimated univariate normalized funding costs, a quantity that we indicate with the symbol $\hat{\Phi}_d$.

To conclude this section, we comment on estimating the MGARCH-VAR model, as postulated in Assumption 4.1. Following an econometric methodology that is well known in the literature on realized volatility (Andersen and Teräsvirta, 2009), as a preliminary estimation step, we replace each of the unobserved Y_d's with its estimates. Hence, we define

$$\hat{Y}_d \doteq \left(\log \hat{\varsigma}^{(1)}_d, \ldots, \log \hat{\varsigma}^{(N)}_d, \log \hat{\Lambda}^{(1)}_d, \ldots, \log \hat{\Lambda}^{(N)}_d, \log \hat{\Phi}_d\right)^\top, \; d = 1, \ldots, D, \tag{26}$$

where $(\hat{\varsigma}^{(1)}_d, \ldots, \hat{\varsigma}^{(N)}_d, \hat{\Lambda}^{(1)}_d, \ldots, \hat{\Lambda}^{(N)}_d, \hat{\Phi}_d)^\top$ is the estimate of $(\varsigma^{(1)}_d, \ldots, \varsigma^{(N)}_d, \Lambda^{(1)}_d, \ldots, \Lambda^{(N)}_d, \Phi_d)^\top$ computed from the intraday data. Then, we fit the parameters of the MGARCH-VAR model conditionally on $\{Y_d\} = \left\{\hat{Y}_d\right\}$.

Despite the constraints discussed in Section 4, the joint estimation of the parameters is still challenging to implement. We consider a multi-step estimation procedure based on a composite likelihood (Pakel et al., 2021) as a possible workaround. Similarly to the traditional quasi-log-likelihood, the composite likelihood is robust to non-gaussian innovations, and, in addition, it typically provides less biased estimates. If the composite likelihood is of bivariate kind, as the one adopted here, the maximization algorithms are numerically stable as they do not require matrix inversions. We refer the reader to Appendix D for all the technical details concerning the estimation of the MGARCH-VAR model.

6 Dataset and preliminary estimation results

We employ a dataset whose constituents are the 250 most liquid stocks (in terms of average transaction volume during the period considered). We have all trades for each stock from January 2006 to December 2014. We apply suitable quality cuts[6], reducing the original

[6]We compute, for each stock and on each day, 1) the total volume traded, 2) the total number of transactions, and 3) the longest time interval with no trading, obtaining three 250×2265 matrices, one for

	Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Average
VAR	$\hat{L}$	5	5	9	4	5	5	5	5	5	5	5	4	5	5	7	5.267
	$\hat{\rho}$	0.996	0.995	0.998	0.996	0.995	0.996	0.995	0.998	0.993	0.997	0.997	0.995	0.997	0.996	0.998	0.996
MGARCH	$\hat{\alpha}_{\varsigma,\varsigma}$	0.017^{*}	0.029^{*}	0.020^{**}	0.019^{***}	0.014^{***}	0.017^{***}	0.019^{***}	0.030^{***}	0.014^{*}	0.074^{***}	0.021^{***}	0.052^{***}	0.025^{***}	0.020^{*}	0.023^{*}	0.026
	$\hat{\alpha}_{\lambda,\varsigma}$	0.016^{**}	0.020^{***}	0.010^{***}	0.022^{***}	0.011^{***}	0.018^{***}	0.020^{***}	0.032^{***}	0.010^{***}	0.025^{***}	0.014^{**}	0.018^{***}	0.020^{***}	0.018^{**}	0.011^{**}	0.018
	$\hat{\alpha}_{\phi,\varsigma}$	0.013^{**}	0.018^{***}	0.015^{*}	0.019^{**}	0.016^{*}	0.011	0.016^{***}	0.025^{***}	0.020^{***}	0.018^{**}	0.018	0.027	0.013^{***}	0.015^{**}	0.015^{*}	0.017
	$\hat{\alpha}_{\lambda,\lambda}$	0.018^{***}	0.038^{**}	0.030^{***}	0.029^{***}	0.023^{***}	0.019^{***}	0.037^{***}	0.039^{***}	0.025^{***}	0.050^{***}	0.032^{***}	0.042^{**}	0.024^{***}	0.042^{**}	0.026^{**}	0.032
	$\hat{\alpha}_{\phi,\lambda}$	0.008	0.003	0.010^{*}	0.013^{*}	0.008	0.011	0.016^{**}	0.011	0.010^{***}	0.012	0.011	0.002	0.014	0.005	0.005	0.009
	$\hat{\alpha}_{\phi,\phi}$	0.027	0.054	0.043	0.054	0.049	0.011	0.034	0.062^{**}	0.051	0.063^{*}	0.031	0.055^{*}	0.041	0.028	0.033	0.043
	$\hat{\alpha}_{\varsigma,\varsigma}+\hat{\beta}_{\varsigma,\varsigma}$	0.998^{***}	0.993^{***}	0.978^{***}	0.999^{***}	0.994^{***}	1.000^{***}	0.995^{***}	0.998^{***}	0.962^{***}	0.854^{***}	0.980^{***}	0.892^{***}	0.998^{***}	0.991^{***}	0.975^{***}	0.974
	$\hat{\alpha}_{\lambda,\varsigma}+\hat{\beta}_{\lambda,\varsigma}$	0.995^{***}	0.961^{***}	0.967^{***}	0.991^{***}	0.983^{***}	0.997^{***}	0.973^{***}	0.986^{***}	0.966^{***}	0.871^{***}	0.968^{***}	0.903^{***}	0.991^{***}	0.966^{***}	0.967^{***}	0.966
	$\hat{\alpha}_{\phi,\varsigma}+\hat{\beta}_{\phi,\varsigma}$	0.978^{***}	0.945^{***}	0.956^{***}	0.955^{***}	0.965^{***}	0.994^{***}	0.950^{***}	0.958^{***}	0.944^{***}	0.850^{***}	0.970^{***}	0.896	0.967^{***}	0.971^{***}	0.952^{***}	0.950
	$\hat{\alpha}_{\lambda,\lambda}+\hat{\beta}_{\lambda,\lambda}$	0.995^{***}	0.958^{***}	0.986^{***}	0.987^{***}	0.986^{***}	0.995^{***}	0.969^{***}	0.980^{***}	0.990^{***}	0.968^{***}	0.982^{***}	0.974^{***}	0.993^{***}	0.968^{***}	0.988^{***}	0.981
	$\hat{\alpha}_{\phi,\lambda}+\hat{\beta}_{\phi,\lambda}$	0.971^{***}	0.908^{***}	0.950^{***}	0.939^{***}	0.948	0.990^{***}	0.930^{***}	0.931^{***}	0.943^{***}	0.914^{***}	0.959^{***}	0.917^{***}	0.966^{***}	0.939^{***}	0.947^{***}	0.944
	$\hat{\alpha}_{\phi,\phi}+\hat{\beta}_{\phi,\phi}$	0.977^{***}	0.947^{***}	0.967^{***}	0.949^{***}	0.967^{***}	0.993^{***}	0.982^{***}	0.963^{***}	0.954^{***}	0.950^{***}	0.977^{***}	0.953^{***}	0.977^{***}	0.980^{***}	0.955^{***}	0.966
	LBQ	4	6	1	8	5	6	5	4	3	2	6	12	7	6	4	5.267
	ARCH	4	3	1	4	1	4	1	4	0	3	1	0	4	2	2	2.267
	$\hat{\kappa}$	11.75	9.50	12.47	13.08	11.39	13.82	12.42	13.42	13.53	13.17	11.51	13.21	12.97	10.19	12.35	12.318

Table 1: MGARCH-VAR estimation results. Summary of the estimation results provided by the MGARCH-VAR model of eqs. (21–25) fitted to the volatility and illiquidity sequences attached to 15 stock samples. Each sample includes ten stocks that deliver 21 volatility and illiquidity sequences in total (10 stocks' volatility sequences, ten stocks' market illiquidity sequences, and the funding illiquidity sequence). $\hat{L}$ is the estimated VAR lag. $\hat{\rho}$ is the spectral radius associated with the companion form of the VAR(L) representation of $\{Y_d\}$. $\hat{\alpha}$ and $\hat{\beta}$ denote the estimated entries of the block MGARCH parameters A and B in equation (24). *LBQ* is the number of sequences of GARCH-standardized residuals, which is significant at 1% level in the Ljung–Box Q-test of no serial correlation. The maximum LBQ is 21, and the expected value under the null hypothesis is 0.21. *ARCH* is the number of sequences of GARCH-standardized residuals, which is significant at 1% level in Engle's test of no *ARCH* effects. The maximum of *ARCH* is 21, and the expected value under the null hypothesis is 0.21. $\hat{\kappa}$ is the estimated kurtosis parameter. Limited to α and $\alpha+\beta$, the significance of the estimated parameters, adjusted for the previous steps of estimation, is reported. The significance levels considered are 5%, 1%, and 1‰, denoted with "$*$", "$**$", and "$***$", respectively. No superscript is reported in the case of insignificance at the 5% level.

dataset to a smaller one made of $N = 150$ stocks and $D = 2265$ trading days. For each day, we obtain transaction prices by sampling (via previous-tick interpolation) on an equispaced deterministic partition at 10-second sampling frequency. This yields, for each day, $T = 2341$ transaction prices for each stock. We begin with some descriptive results, summarized in Figures 2-3.

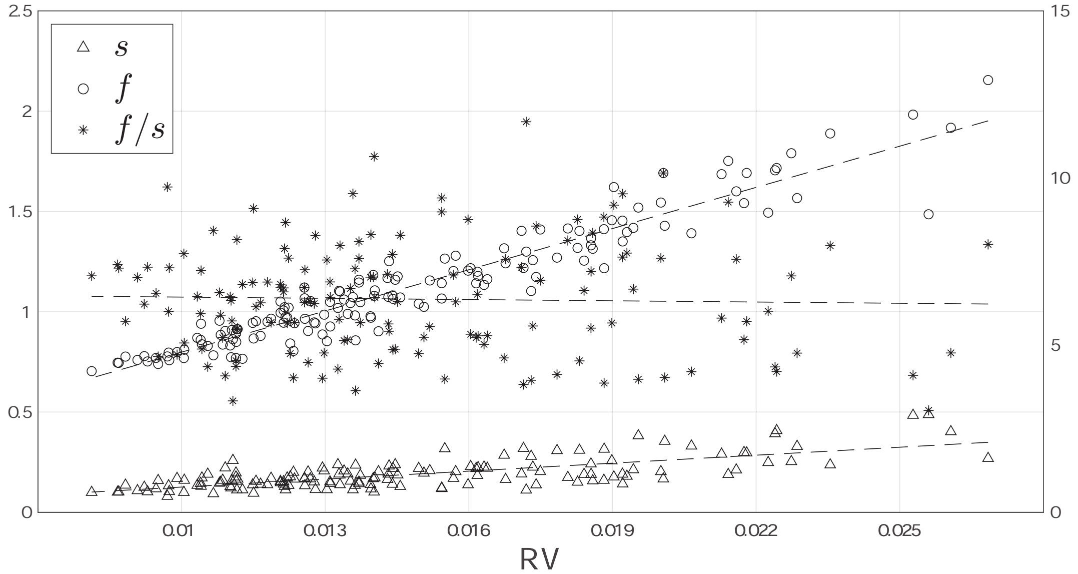

Figure 2: We report, for each stock in the data sample and as a function of the stocks' realized volatilities (the square roots of the realized variances), the estimated half bid-ask spread (empty triangles), the estimated funding cost (empty circles), and their ratio (black stars). For each variable, we also show the corresponding regression line. The slope is not significantly different from zero only in the case of the funding cost to bid-ask spread ratio. The left vertical axis shows the range of the first two variables (in basis points), while the right vertical axis shows the range of the funding cost to bid-ask spread ratio, which is a pure number.

Figure 2 shows, for each stock in the dataset and as a function of the stocks' realized volatilities, the estimated half bid-ask spread (in the plot indicated as s), the estimated funding cost (marked as f) and their ratio f/s. While both s and f appear to be highly dependent on the stock's realized volatility (as witnessed by the slope of the regression

each characteristic. We flag all the daily entries of the three matrices with a total log-volume smaller than 12.5, a total number of transactions smaller than 500, or a maximum time length of no trading larger than ten minutes. After removing all flagged days, we keep only stocks with daily returns over at least 97% of the sample.

line, which turns out to be statistically different from zero), their ratio is not, with a regression line statistically indistinguishable from a horizontal line. This empirical evidence supports the approximation imposed by equation (19), which translates the assumption of a common, across-stocks value of the funding cost to bid-ask spread ratio.

Figure 3 reports the daily average (across-stocks) market illiquidity and the daily funding illiquidity estimates, along with the TED spread and the timeline of the Global Financial Crisis started in 2007. Notably, the correlation of the estimated funding illiquidity with the TED spread is 62%. Furthermore, compared with the TED spread, the estimated funding illiquidity is more sensitive to macroeconomic announcements, such as the rise of the U.S. unemployment rate in October 2009.

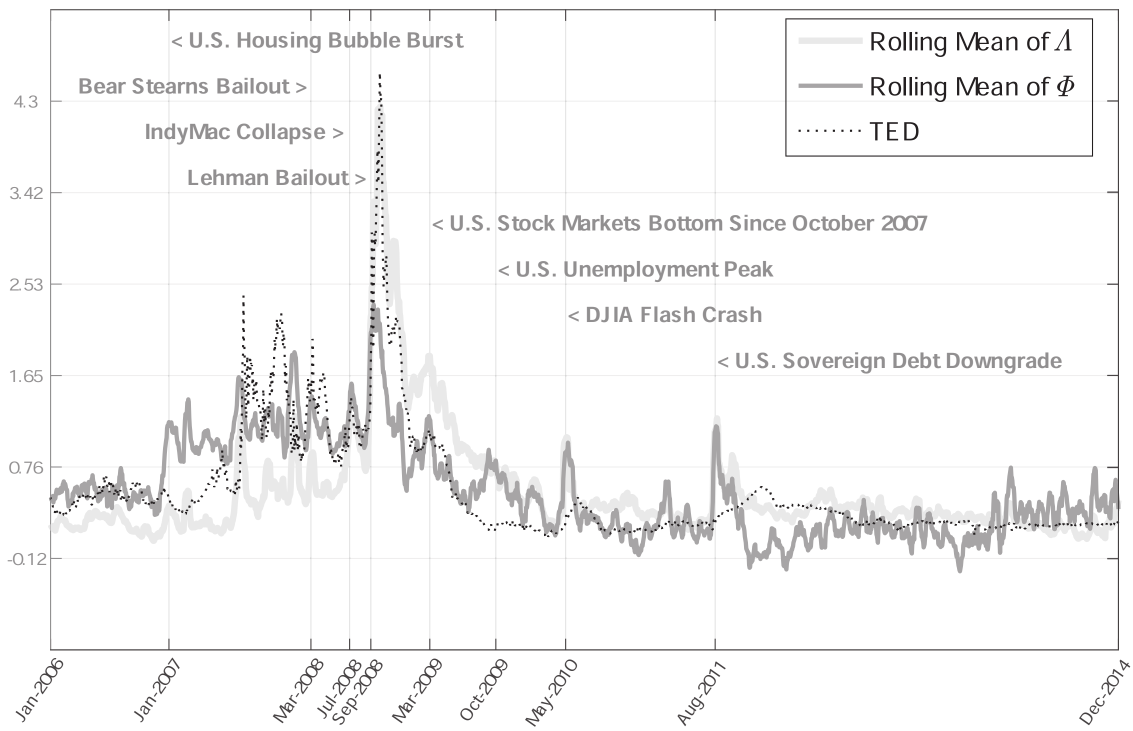

Figure 3: We report, along with the timeline of the Global Financial Crisis started in 2007, the daily estimated average (across-stocks) market illiquidity (light gray thin line), the daily funding illiquidity (dark gray thick line), and the TED spread (dotted line). The time series of average market illiquidity and funding illiquidity are reported as moving averages over a window of eleven days of their daily estimates, normalized to have the same mean and standard deviation of the TED spread.

Rather than fitting the MGARCH-VAR model once and for all to the 150 stocks, we split the dataset into 15 samples of $N = 10$ stocks each. In this way, the stability of the estimation output across different stock samples is assessable. Stocks' heterogeneity within each sample is ensured by drawing the ten stocks from a different volatility decile each. Sampling is done without repetition so that all available 150 stocks are selected.

Table 1 reports the MGARCH-VAR estimates from the 15 samples. The estimation output is overall stable across samples. The estimated VAR lag $\hat{L}$ (defined in equation (21) of Assumption 4.1) is equal to 5.267 days on average (about one week of daily sessions). The spectral radius, $\hat{\rho}$, associated with the companion form of the VAR equation, is close to one, suggesting that the volatility and illiquidity processes are both highly persistent. Also, the persistence of the variance-covariance processes, estimated by $\hat{\alpha} + \hat{\beta}$, is close to one.

As a first specification test, denoted with *LBQ*, we compute the number of sequences of GARCH-standardized residuals, which is significant at 1% level in the Ljung–Box Q-test of no serial correlation. For a sample of $M = 21$ sequences, the maximum *LBQ* is 21. Under the null hypothesis, the expected value is $21 \cdot 0.01 = 0.21$. We get 5.6 rejections on average, a significant value that we still consider acceptable given the common lag across the univariate sequences.

As a second specification test, denoted with *ARCH*, we compute the number of sequences of GARCH-standardized residuals which is significant at 1% level in Engle's test of no *ARCH* effects[7]. Rejecting the null hypothesis of no *ARCH* effects would provide evidence of misspecified conditional second-moment dynamics. Similarly to *LBQ*, the maximum value of the test statistic is 21, and the expected value under the null hypothesis is 0.21. We get about two rejections on average, an acceptable outcome.

7 New empirical insights on volatility and illiquidity dynamics

The mutual interaction between market and funding illiquidity is expected to be multifaceted and non-trivial. Our framework is designed to provide daily realized measures of these market features. Still, it works in the spirit of "let the data speak" concerning their reciprocal interaction as daily time series. Consequently, we take insights from the extant theoretical literature to guide our empirical investigation on the mutual influence between

[7]The lag parameter in both *LBQ* and *ARCH* test is set to the integer part of log(2265), where 2265 is the length of the sequence.

market and funding illiquidity. In particular, following the results of Brunnermeier and Pedersen (2009), we test four conjectures on the dataset described at the beginning of Section 6. These are formulated as claims on the impact that market and funding illiquidity should have on each other. Before entering technical details, we formulate the four conjectures in the following list.

(i) *Volatility and illiquidity co-movements*. Stocks' fundamental volatilities, stocks' market illiquidities, and funding illiquidity co-move.

(ii) *Asymmetric co-movements*. Volatility and illiquidity co-movements are stronger when funding liquidity tightens.

(iii) *Quality-and-liquidity*. Stocks with lower volatility are characterized by higher market liquidity.

(iv) *Flight-to-quality*. When funding liquidity is low on average, high-volatility stocks are more sensitive to changes in funding liquidity than low-volatility stocks.

Given their general formulation, testing these four conjectures is far from obvious. The test should consist of a statistical procedure capable of providing reliable evidence in favor or against the claim that is put forward. Here, we proceed in this way.

Let us denote with $\mathcal{M} \doteq \{f^{(Y)}(\mathcal{Y};\boldsymbol{\theta}); \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ the MGARCH-VAR model, introduced in Section 4, for the vector time series $\{Y_d\}$, defined in equation (20), where we have used the notation $\mathcal{Y} \doteq \mathcal{Y}_D = \{Y_D, Y_{D-1}, \ldots, Y_1\}$, $\boldsymbol{\theta} \doteq \{L, C, W, A, B, \kappa\}$, and $C \doteq \{c_0, C_1, \ldots, C_L\}$. The null hypothesis, $\mathcal{H}_0$, is that the density of $\mathcal{Y}$ is in $\mathcal{M}$ (correctly specified model). The alternative hypothesis, $\mathcal{H}_1$, is that the density of $\mathcal{Y}$ is not in $\mathcal{M}$ (misspecified model).

To each of the four conjectures, we associate a statistic, say τ, that represents the market feature put forward by the conjecture itself. As an example, concerning the claim of the *Volatility and illiquidity co-movements* conjecture, the statistic τ ought to be any observable deemed to represent co-movements between volatility and illiquidity (market and funding). More generally, regardless of the claim being tested, the statistic τ must be a function of the whole path $\{Y_D, \ldots, Y_1\}$, i.e. $\tau \doteq \tau(\mathcal{Y})$. In all of our applications, we use a real-valued statistic; that is, we assume that $\tau \in \mathbb{R}$ (in other words, we do not consider vector-valued statistics).

Under $\mathcal{H}_0$, the cumulative distribution function (CDF) of τ is denoted with

$$F_0^{(\tau)}(u) \doteq \int_{\mathcal{Y}:\tau(\mathcal{Y})\leq u} f^{(Y)}(\mathcal{Y};\boldsymbol{\theta}_0)d\mathcal{Y}, \quad u \in \mathbb{R}, \tag{27}$$

where $\boldsymbol{\theta}_0$ is the true parameter value.

Testing $\mathcal{H}_0$ against $\mathcal{H}_1$, via τ as a test statistic, requires to know $\mathcal{Y}$ and the functional form of the distribution $u \to F_0^{(\tau)}(u)$. Since they are both unknown (except for a few cases in which the distribution of τ can be computed analytically), we replace $\mathcal{Y}$ with its realized counterpart $\hat{\mathcal{Y}} \doteq \left\{\hat{Y}_D, \hat{Y}_{D-1}, \ldots, \hat{Y}_1\right\}$, obtained from equation (26), while we estimate the distribution $F_0^{(\tau)}$ via parametric bootstrap[8] (Efron and Tibshirani, 1993; MacKinnon, 2009). In what follows, we use the notation $\hat{\tau} \doteq \tau\left(\hat{\mathcal{Y}}\right)$ to denote the statistic τ computed from $\hat{\mathcal{Y}}$.

If the observed statistic $\hat{\tau}$ is compatible (in a statistical sense) with the distribution under the null, i.e., if $\mathcal{H}_0$ is *not* rejected, then we conclude that the volatility and illiquidity dynamics is of MGARCH-VAR type. Conditionally on this finding, we test the statistical significance of the bootstrap expected value of the statistic τ under the null. Should the expected value be statistically different from zero, we could have evidence in favor or against the conjecture, depending on the statistic and its sign. Later in the text, we go into more detail about this feature.

Should, instead, $\mathcal{H}_0$ be rejected, we interpret this event as empirical evidence of the presence of *extra* volatility and illiquidity dynamic, i.e., in addition to the volatility and illiquidity dynamic captured by the MGARCH-VAR model. Again, depending on the statistic and its sign, this rejection could be in favor or against the conjecture under study, as will be apparent from the following four sections.

7.1 Volatility and illiquidity co-movements

Let (i, j) be a couple of given indexes, each from 1 to M, where M is the number of elements of the vectors $\left\{\hat{Y}_d\right\}$ defined in (26). Let $\hat{W}^{(i,j)}$ be the estimated sample second moment of the residuals $\left(\hat{U}_d^{(i)}, \hat{U}_d^{(j)}\right)$ of the MGARCH-VAR model in equation (21). Let $\mathsf{HAC}[\,\cdot\,]$ denote heteroskedasticity and autocorrelation robust standard errors[9] and let $t_{i,j} \doteq \hat{W}^{(i,j)} \big/ \mathsf{HAC}\left[\hat{W}^{(i,j)}\right]$. Consider the statistic

$$\mathsf{CM} \doteq \sum_{i<j=2,\ldots,M} \mathbf{1}(t_{i,j} > 2.33) - \sum_{i<j=2,\ldots,M} \mathbf{1}(t_{i,j} < -2.33). \tag{28}$$

[8]The parametric bootstrap requires to draw K sequences from the DGP of the MGARCH-VAR model, indexed by the estimated parameters. From each simulated sequence, a replicate of the statistic τ, denoted with $\tilde{\tau}_j$, for $j = 1, \ldots, K$, is computed. The empirical CDF of the collection $\{\tilde{\tau}_j\}$, denoted with $\hat{F}$, is then taken as the estimate of the true distribution $F_0^{(\tau)}$, defined in equation (27). Under appropriate conditions and under the null $\mathcal{H}_0$, it can be proved that $\hat{F}$ is a consistent estimator of $F_0^{(\tau)}$.

[9]In this paper we adopt the formula proposed by Newey and West (1987).

The CM statistic represents thus the number of sample covariances of $\{\hat{U}_d\}$ that are significant at 1% level in a right-tailed t-test, in excess with respect to the number of sample covariances of $\{\hat{U}_d\}$, significant at 1% level in a left-tailed t-test. The rationale behind the quantity in (28) is simple: assuming a diagonal W (i.e., in the absence of co-movements in the MGARCH-VAR residuals), in the limit $D \longrightarrow \infty$, the distribution of CM is centered at zero. Hence, we interpret statistically positive (resp. negative) values of CM as a signal of positive (resp. negative) co-movements in the VAR residuals[10]. Accordingly, being CM a measure of volatility and illiquidity co-movements, we select it as a statistic associated with the first conjecture.

Concerning the 15 samples of 10 stocks each, the test results are reported in Table 2.

Consider, for instance, sample 12. The value of the CM statistic is 196, which is insignificant, at 5% level, under $\mathcal{H}_0$. This does not mean we do not find evidence favoring the first conjecture. In fact, for sample 12, the bootstrap expected value under $\mathcal{H}_0$ for the statistic CM is equal to 187.2 and, most importantly, is significant at 1‰. Therefore, conditionally on the assumption of a correctly specified model, the hypothesis of zero MGARCH-VAR co-movements is rejected at 1‰ in favor of strictly positive co-movement.

In some other cases, the statistic is more sensitive. Consider, for this purpose, sample 1. In this case, the CM shows a value of 160, which turns out to be significant at 1%. Similarly, its expected value under the null $\mathcal{H}_0$ is 112.8, which is significant at 1‰. As anticipated in the previous section, we view this finding as supporting the existence of extra *positive* co-movements, i.e., in addition to the positive co-movements captured by the MGARCH-VAR model. This empirical result is stable across all samples, except for sample 9 and sample 15, for which the null hypothesis of MGARCH-VAR dynamics is not rejected.

It's worth investigating the dynamics of co-movements more thoroughly. Following indications from the theoretical and empirical literature[11], we look for the presence of

[10] Our measure of co-movements, CM, is inspired by Proposition 6-(i) of Brunnermeier and Pedersen (2009). Here, co-movements are defined in terms of conditional covariances of $\{Y_d\}$. Within our MGARCH-VAR framework, the conditional covariances of $\{Y_d\}$ coincide with the conditional covariances of the VAR residuals. The conditional covariances of the VAR residuals are the main ingredient for the computation of CM. Of course, other measures of co-movements would serve the purpose.

[11] The equilibrium model of Brunnermeier and Pedersen (2009) is characterized by commonality in liquidities. Chordia et al. (2000) and Karolyi et al. (2012) provide an empirical study of the commonality in liquidities.

co-movements in illiquidity. For this purpose, we partition the matrix $\hat{W}$ as

$$\hat{W} = \begin{pmatrix} \hat{W}_{\varsigma,\varsigma}\mathbb{I}_{N,N} & \hat{W}_{\varsigma,\lambda}\mathbb{I}_{N,N} & \hat{W}_{\varsigma,\phi}\mathbb{I}_{N,1} \\ \hat{W}_{\lambda,\varsigma}\mathbb{I}_{N,N} & \hat{W}_{\lambda,\lambda}\mathbb{I}_{N,N} & \hat{W}_{\lambda,\phi}\mathbb{I}_{N,1} \\ \hat{W}_{\phi,\varsigma}\mathbb{I}_{1,N} & \hat{W}_{\phi,\lambda}\mathbb{I}_{1,N} & \hat{W}_{\phi,\phi}\mathbb{I}_{1,1} \end{pmatrix},$$

with obvious notation. From the blocks in $\hat{W}$ we compute test statistics, analogous to CM, denoted with $\mathsf{CM}_{\varsigma,\varsigma}$, $\mathsf{CM}_{\lambda,\varsigma}$, $\mathsf{CM}_{\phi,\varsigma}$, $\mathsf{CM}_{\lambda,\lambda}$, and $\mathsf{CM}_{\phi,\lambda}$.[12] The corresponding empirical findings are reported in the rows from the second to the sixth of Table 2. We find evidence of extra positive co-movements in stocks' market illiquidities (see row $\mathsf{CM}_{\lambda,\lambda}$) and strong evidence of extra positive co-movements in stocks' fundamental volatilities (see row $\mathsf{CM}_{\varsigma,\varsigma}$). We notice that, even in the absence of extra positive co-movements, the MGARCH-VAR co-movements are of a positive kind (see row $\mathsf{CM}_{\phi,\varsigma}$).

7.2 Asymmetric co-movements

We say that volatility and illiquidity co-movements are asymmetric if they are stronger when the funding liquidity tightens[13]. A way to look for such a kind of asymmetry is to regress the pairwise conditional correlations of $\left\{\hat{Y}_d\right\}$ on the (log) funding illiquidity. To be more specific, let us consider the regressions

$$\hat{\varrho}_d^{(i,j)} = a_0^{(i,j)} + a_1^{(i,j)} \log \hat{\Phi}_d + \varepsilon_d^{(i,j)},$$

for $i < j$, where

$$\hat{\varrho}_d^{(i,j)} \doteq \hat{H}_d^{(i,j)} \Big/ \sqrt{\hat{H}_d^{(i,i)} \cdot \hat{H}_d^{(j,j)}} \tag{29}$$

is the estimated conditional correlation of $(Y_d^{(i)}, Y_d^{(j)})$ delivered by the MGARCH-VAR estimation output. If $a_1^{(i,j)}$ is positive, the estimated co-movements in $(Y_d^{(i)}, Y_d^{(j)})$ increase with the funding illiquidity. Our across-stocks measure of asymmetric co-movements is thus defined as

$$\mathsf{AC} \doteq \sum_{i<j=2,\dots,M} \mathbf{1}\left(t_{i,j} > 2.33\right) - \sum_{i<j=2,\dots,M} \mathbf{1}\left(t_{i,j} < -2.33\right),$$

where now the t-statistics are defined as $t_{i,j} \doteq \hat{a}_1^{(i,j)}/\mathsf{HAC}[\hat{a}_1^{(i,j)}]$, with $\hat{a}_1^{(i,j)}$ the OLS estimator of $a_1^{(i,j)}$. Positive (resp. negative) values of AC indicate positive (resp. negative) asymmetric

[12] For instance, in $\mathsf{CM}_{\lambda,\varsigma} = \sum_{N+1\le i\le 2N, 1\le j\le N} \mathbf{1}(t_{i,j} > 2.33) - \sum_{N+1\le i\le 2N, 1\le j\le N} \mathbf{1}(t_{i,j} < -2.33)$, only the pairwise second moments between a stock market illiquidity and stock volatility are considered.

[13] See, e.g., Brunnermeier and Pedersen (2009), Proposition 6-(ii) and their Section 6.

Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Range	*
CM	$160^{**}_{112.8^{***}}$	$154^{*}_{126.7^{***}}$	$184^{*}_{167.1^{***}}$	$195^{**}_{148.7^{***}}$	$190^{*}_{171.9^{***}}$	$164^{*}_{101.2^{***}}$	$195^{**}_{173.5^{***}}$	$165^{*}_{130.6^{***}}$	$187_{172.0^{***}}$	$197^{*}_{185.2^{***}}$	$182^{*}_{163.0^{***}}$	$196_{187.2^{***}}$	$173^{*}_{135.3^{***}}$	$147^{*}_{128.5^{***}}$	$190_{176.0^{***}}$	±210	12
$\mathsf{CM}_{\varsigma,\varsigma}$	$45^{***}_{31.6^{***}}$	$42^{*}_{32.1^{***}}$	$45^{***}_{44.6^{***}}$	$45^{***}_{31.8^{***}}$	$45^{***}_{44.2^{***}}$	$44^{*}_{21.4^{***}}$	$45^{***}_{41.7^{***}}$	$44^{*}_{29.6^{***}}$	$45^{***}_{45.0^{***}}$	$45^{***}_{45.0^{***}}$	$45^{***}_{44.1^{***}}$	$45^{***}_{45.0^{***}}$	$44_{34.4^{***}}$	$42_{38.2^{***}}$	$45^{***}_{44.9^{***}}$	±45	13
$\mathsf{CM}_{\lambda,\varsigma}$	$61_{41.9^{***}}$	$51_{46.2^{***}}$	$87_{79.9^{***}}$	$92^{**}_{68.4^{***}}$	$88_{78.9^{***}}$	$66_{41.8^{***}}$	$90^{*}_{77.9^{***}}$	$77^{*}_{61.3^{***}}$	$89_{84.2^{***}}$	$98^{*}_{92.0^{***}}$	$82_{72.8^{***}}$	$93_{91.3^{***}}$	$79_{61.9^{***}}$	$59_{50.6^{***}}$	$92_{86.2^{***}}$	±100	4
$\mathsf{CM}_{\phi,\varsigma}$	$7_{5.8^{***}}$	$8_{5.7^{***}}$	$3_{3.6^{***}}$	$5_{5.6^{***}}$	$5_{5.2^{***}}$	$6_{5.6^{***}}$	$6_{6.0^{***}}$	$3_{4.1^{***}}$	$7_{6.5^{***}}$	$3_{4.8^{***}}$	$6_{5.6^{***}}$	$7_{6.2^{***}}$	$6_{5.8^{***}}$	$7_{5.1^{***}}$	$4_{5.2^{***}}$	±10	0
$\mathsf{CM}_{\lambda,\lambda}$	$37^{*}_{24.0^{***}}$	$43^{**}_{33.2^{***}}$	$42^{**}_{31.5^{***}}$	$44^{**}_{34.2^{***}}$	$43^{**}_{34.7^{***}}$	$39^{**}_{24.3^{***}}$	$44^{*}_{38.2^{***}}$	$31_{27.0^{***}}$	$38^{*}_{28.1^{***}}$	$42^{*}_{35.0^{***}}$	$39^{*}_{31.2^{***}}$	$42^{*}_{35.8^{***}}$	$35_{24.7^{***}}$	$30_{25.6^{***}}$	$42^{**}_{31.6^{***}}$	±45	12
$\mathsf{CM}_{\phi,\lambda}$	$10^{***}_{9.5^{***}}$	$10^{***}_{9.5^{***}}$	$7_{7.4^{***}}$	$9_{8.8^{***}}$	$9_{8.9^{***}}$	$9_{8.1^{***}}$	$10^{***}_{9.7^{***}}$	$10^{***}_{8.7^{***}}$	$8_{8.2^{***}}$	$9^{*}_{8.5^{***}}$	$10^{***}_{9.3^{***}}$	$9_{8.9^{***}}$	$9_{8.6^{***}}$	$9_{9.0^{***}}$	$7_{8.1^{***}}$	±10	6
AC	$81^{***}_{-0.1}$	$123^{***}_{-0.0}$	$62^{*}_{0.0}$	$89^{**}_{-0.4}$	$117^{***}_{0.1}$	$44_{0.1}$	$127^{***}_{-0.0}$	$80^{**}_{-0.1}$	$57^{*}_{-0.0}$	$43^{**}_{0.1}$	$132^{***}_{-0.1}$	$37^{*}_{0.1}$	$104^{***}_{0.3}$	$54^{**}_{0.1}$	$93^{***}_{-0.0}$	±210	14
$\mathsf{AC}_{\varsigma,\varsigma}$	$28^{**}_{-0.1}$	$36^{***}_{-0.0}$	$39^{***}_{0.0}$	$27^{*}_{-0.2}$	$36^{***}_{0.0}$	$7_{0.0}$	$42^{***}_{0.1}$	$23^{*}_{0.0}$	$41^{***}_{0.1}$	$44^{***}_{0.1}$	$44^{***}_{0.1}$	$35^{***}_{0.0}$	$32^{**}_{0.0}$	$21^{*}_{-0.0}$	$43^{***}_{-0.1}$	±45	14
$\mathsf{AC}_{\lambda,\varsigma}$	$31^{*}_{0.0}$	$54^{***}_{0.0}$	$-5_{0.1}$	$27_{-0.2}$	$46^{**}_{0.0}$	$12_{0.1}$	$38^{**}_{-0.1}$	$32^{*}_{-0.1}$	$-3_{-0.1}$	$-28^{**}_{0.0}$	$45^{***}_{-0.0}$	$-22^{*}_{-0.0}$	$50^{**}_{0.2}$	$7_{0.1}$	$13_{0.0}$	±100	9
$\mathsf{AC}_{\phi,\varsigma}$	$8^{**}_{0.0}$	$9^{***}_{-0.0}$	$9^{***}_{-0.0}$	$7^{**}_{-0.0}$	$9^{***}_{-0.0}$	$7^{*}_{0.0}$	$9^{***}_{-0.0}$	$7^{**}_{-0.0}$	$10^{***}_{-0.0}$	$10^{***}_{-0.0}$	$9^{***}_{-0.0}$	$10^{***}_{0.0}$	$9^{***}_{0.0}$	$6^{*}_{0.1^{*}}$	$9^{**}_{0.0}$	±10	15
$\mathsf{AC}_{\lambda,\lambda}$	$16^{*}_{-0.1}$	$20^{***}_{0.0}$	$20^{**}_{-0.1}$	$27^{***}_{0.0}$	$25^{***}_{-0.0}$	$17^{*}_{-0.0}$	$29^{***}_{-0.1}$	$14^{*}_{0.1}$	$7_{0.0}$	$15^{*}_{0.0}$	$31^{***}_{-0.1}$	$15^{**}_{0.1}$	$13_{0.1}$	$15^{**}_{-0.0}$	$30^{***}_{-0.1}$	±45	13
$\mathsf{AC}_{\phi,\lambda}$	$-2_{0.0}$	$4_{-0.1}$	$-1_{0.0}$	$1_{0.0}$	$1_{0.0}$	$1_{0.0}$	$9^{***}_{0.0}$	$4_{-0.0}$	$2_{0.0}$	$2_{-0.0}$	$3_{0.0}$	$-1_{-0.0}$	$0_{-0.0}$	$5_{0.0}$	$-2_{0.1^{*}}$	±10	1
QL_1	$0.45^{*}_{0.34^{***}}$	$0.48^{***}_{0.33^{***}}$	$0.60^{**}_{0.55^{***}}$	$0.51^{***}_{0.34^{***}}$	$0.70_{0.71^{***}}$	$0.46^{*}_{0.34^{***}}$	$0.67^{*}_{0.61^{***}}$	$0.35^{***}_{0.23^{***}}$	$0.53_{0.56^{***}}$	$0.61_{0.62^{***}}$	$0.53^{*}_{0.44^{***}}$	$0.67_{0.65^{***}}$	$0.48_{0.43^{***}}$	$0.27_{0.17^{***}}$	$0.42_{0.41^{***}}$	±1	8
QL_2	$6^{**}_{6.0^{***}}$	$8_{6.7^{***}}$	$9_{8.4^{***}}$	$8^{*}_{7.9^{***}}$	$10^{***}_{10.0^{***}}$	$8_{7.7^{***}}$	$10^{***}_{9.6^{***}}$	$7_{7.0^{***}}$	$10^{***}_{9.9^{***}}$	$10^{***}_{10.0^{***}}$	$10^{***}_{9.2^{***}}$	$10^{***}_{10.0^{***}}$	$8^{*}_{8.0^{***}}$	$6_{5.3^{***}}$	$10^{***}_{9.8^{***}}$	±10	10
FQ_1	$-0.26_{-0.24^{***}}$	$-0.10_{-0.08^{***}}$	$0.02_{0.02^{***}}$	$-0.14^{*}_{-0.07^{***}}$	$-0.24_{-0.17^{***}}$	$-0.17_{-0.10^{***}}$	$0.09_{0.11^{***}}$	$-0.34^{*}_{-0.27^{***}}$	$-0.39_{-0.36^{***}}$	$-0.29_{-0.26^{***}}$	$-0.34^{**}_{-0.25^{***}}$	$-0.16_{-0.15^{***}}$	$0.02_{0.10^{***}}$	$-0.12^{***}_{-0.04^{***}}$	$-0.28_{-0.24^{***}}$	±1	4
FQ_2	$5^{**}_{-0.2^{***}}$	$8^{***}_{-0.8^{***}}$	$9^{***}_{-0.1^{***}}$	$5^{*}_{-0.3^{***}}$	$4^{*}_{-0.6^{***}}$	$8^{***}_{-0.5^{***}}$	$4^{*}_{-0.6^{***}}$	$8^{***}_{-0.5^{***}}$	$4^{**}_{-0.2^{***}}$	$8^{***}_{-0.5^{***}}$	$5^{**}_{-0.4^{***}}$	$8^{**}_{-0.6^{***}}$	$7^{**}_{-0.5^{***}}$	$6^{***}_{-0.4^{***}}$	$5^{**}_{-0.3^{***}}$	±10	15

Table 2: Test results. For each test in row, and each sample in column, the observed test statistic, $\hat{\tau}$, and the mean of the bootstrap replicates, $\tilde{\tau}$, are reported, arranged as $\hat{\tau}^{sgnf}_{\tilde{\tau}^{sgnf}}$. For $\hat{\tau}$, superscript "*sgnf*" denotes significance in the bootstrap test. For $\tilde{\tau}$, superscript "*sgnf*" denotes significance of the attached *t*-statistic in the bootstrap experiment. The levels considered are 5%, 1%, and 1‰, denoted with "*", "**", and "***", respectively. No superscript is reported in the case of insignificance at the 5% level. The second-last column reports the range of the test statistics. The last column reports the number of significant samples in the bootstrap test. Consider, for instance, the result of CM on sample 1, written as $160^{**}_{112.8^{***}}$. The observed CM equals 160, which is significant in the bootstrap test at the 1% level. The bootstrap mean of CM is 112.8, which is significant in the bootstrap experiment at 1‰ level. The range of CM is ±210, and the number of samples that are significant in the bootstrap test is 12.

co-movements in the components of $\{\hat{Y}_d\}$. The theoretical predictions are typically in favor of positive asymmetry.

Moving to the test results in Table 2, we notice that, apart from sample 6, the null hypothesis of MGARCH-VAR dynamics is always rejected. Since the observed AC is always positive, the evidence is in favor of positively asymmetric co-movements. We also notice that the bootstrap mean of AC is never significant in the bootstrap experiment. This suggests that the co-movements generated by the MGARCH-VAR process are symmetric.

The equilibrium model of Brunnermeier and Pedersen (2009) is characterized by "commonality of fragility", a prediction of positively asymmetric co-movements in stocks' market illiquidities[14]. In our framework, a test statistic specific for commonality of fragility can be computed similarly to AC, as long as only the conditional correlations between stocks' market illiquidities are involved. The test statistic, denoted with $\mathsf{AC}_{\lambda,\lambda}$, reads

$$\mathsf{AC}_{\lambda,\lambda} \doteq \sum_{N<i<j=N+2,\dots,2N} \mathbf{1}\left(t_{i,j} > 2.33\right) - \sum_{N<i<j=N+2,\dots,2N} \mathbf{1}\left(t_{i,j} < -2.33\right).$$

In line with the prediction, the hypothesis of symmetric (MGARCH-VAR) co-movements is rejected on 13 samples (see row $\mathsf{AC}_{\lambda,\lambda}$ of Table (2)).

Also of interest are the tests denoted by $\mathsf{AC}_{\varsigma,\varsigma}$, $\mathsf{AC}_{\lambda,\varsigma}$, $\mathsf{AC}_{\phi,\varsigma}$, and $\mathsf{AC}_{\phi,\lambda}$, with obvious notation. As shown in Table 2, most tests are in favor of the prediction of positively asymmetric co-movements. One exception is the test based on $\mathsf{AC}_{\phi,\lambda}$, according to which, apart from sample 7, the null hypothesis of symmetric (MGARCH-VAR) co-movements is never rejected. We conclude that the level of funding illiquidity does not affect the co-movements between funding illiquidity and stock market illiquidity.

7.3 Quality-and-liquidity

The quality-and-liquidity conjecture predicts that stocks with lower volatility should be associated with higher market liquidity[15]. We test for quality and liquidity both at a cross-sectional level and at a time series level.

At a cross-sectional level, we measure quality-and-liquidity with

$$\mathsf{QL}_1 \doteq \frac{1}{D} \sum_{d=1,\dots,D} \hat{\varrho}_d^{(\mathsf{QL})},$$

[14] See Proposition 6-(ii) in Brunnermeier and Pedersen (2009).

[15] See, e.g., Brunnermeier and Pedersen (2009), Proposition 6-(iii).

where $\hat{\varrho}_d^{(\mathsf{QL})}$ is the sample rank correlation between the estimated stocks' fundamental volatilities, $(\hat{\varsigma}_d^{(1)}, \ldots, \hat{\varsigma}_d^{(N)})$, and the estimated stocks' market illiquidities, $(\hat{\Lambda}_d^{(1)}, \ldots, \hat{\Lambda}_d^{(N)})$. Positive values of QL_1 indicate cross-sectional quality-and-liquidity.

The tests based on QL_1 return evidence of extra (with respect to the MGARCH-VAR dynamics) cross-sectional quality-and-liquidity on eight samples (see Table 2). In the remaining seven samples, the null hypothesis of MGARCH-VAR dynamics is not rejected. In these samples, the bootstrap mean of QL_1 is positive and highly significant, which shows that the MGARCH-VAR model captures significant cross-sectional quality and liquidity.

At a time series level, we regard quality-and-liquidity as the propensity of stocks to become less liquid when they become more volatile (a sort of *dynamic* quality-and-liquidity effect). Given the regression model

$$\log \hat{\Lambda}_d^{(i)} = b_0^{(i)} + b_1^{(i)} \log \hat{\varsigma}_d^{(i)} + \varepsilon_{\mathsf{QL},d}^{(i)},$$

a measure of dynamic quality-and-liquidity for stock i is the slope of the regression, $b_1^{(i)}$. If $b_1^{(i)} > 0$, there is dynamic quality-and-liquidity in stock i. Therefore, as an overall measure, across stocks, of dynamic quality-and-liquidity we can compute

$$\mathsf{QL}_2 \doteq \sum_{i=1,\ldots,N} \mathbf{1}(t_i > 2.33) - \sum_{i=1,\ldots,N} \mathbf{1}(t_i < -2.33),$$

with t-statistics defined as $t_i \doteq \hat{b}_1^{(i)} / \mathsf{HAC}[\hat{b}_1^{(i)}]$, where $\hat{b}_1^{(i)}$ is the OLS estimator of $b_1^{(i)}$. Large values of QL_2 indicate overall, across stocks, dynamic quality-and-liquidity[16].

We find evidence of extra dynamic quality and liquidity on ten samples (see Table 2). The null hypothesis of MGARCH-VAR dynamics in the remaining five samples is not rejected. In these samples, the bootstrap mean of QL_2 is positive and highly significant, which means that the MGARCH-VAR model captures significant dynamic quality and liquidity.

7.4 Flight-to-quality

According to the predictions of the flight-to-quality conjecture, we expect that when funding liquidity is tight, high-volatility stocks are more sensitive to changes in the funding liquidity than low-volatility stocks[17].

[16] In Section 7.1 we tested for co-movements in the shocks to $\left\{\Lambda_d^{(i)}\right\}$ and $\left\{\varsigma_d^{(i)}\right\}$. The test based on QL_2 can be seen as a test of co-movements in the levels of $\left\{\Lambda_d^{(i)}\right\}$ and $\left\{\varsigma_d^{(i)}\right\}$.

[17] See, e.g., Brunnermeier and Pedersen (2009), Prop. 6-(iv), equation (30).

As a measure of *cross-sectional* flight-to-quality we compute

$$\mathsf{FQ}_1 \doteq \frac{1}{D} \sum_{d=1,\ldots,D} \hat{\varrho}_d^{(\mathsf{FQ})},$$

where $\hat{\varrho}_d^{(\mathsf{FQ})}$ is the sample rank correlation between the estimated stocks' fundamental volatilities, $(\hat{\varsigma}_d^{(1)}, \ldots, \hat{\varsigma}_d^{(N)})$, and the vector $(\hat{\varrho}_d^{(N+1,M)}, \ldots, \hat{\varrho}_d^{(2N,M)})$ of the estimated conditional correlations of $(\log \hat{\Lambda}_d^{(i)}, \log \hat{\Phi}_d)$, for $i = 1, \ldots, N$ (see equation (29)). Positive values of FQ_1 are associated with flight-to-quality.

Looking at the empirical results in Table 2, we notice that the null hypothesis of MGARCH-VAR dynamics is *not* rejected on 11 samples. Surprisingly, the bootstrap mean of FQ_1 is overall negative and significant, thereby suggesting a sort of flight-*from*-quality captured by the MGARCH-VAR model. The unexpected result is likely due to model misspecification. The misspecification is not detected by FQ_1, perhaps because of poor power.

At a time series level, a measure of flight-to-quality for stock i is the interaction coefficient in the regression model

$$\log \hat{\Lambda}_d^{(i)} = c_0^{(i)} + c_1^{(i)} \log \hat{\varsigma}^{(i)} + c_2^{(i)} \log \hat{\Phi}_d + c_3^{(i)} \left(\hat{\varsigma}_d^{(i)} \log \hat{\Phi}_d \right) + \varepsilon_{\mathsf{FQ},d}^{(i)}.$$

If $c_3^{(i)}$ is positive, $\left\{ \hat{\Lambda}_d^{(i)} \right\}$ is more sensitive to $\left\{ \hat{\Phi}_d \right\}$ in periods of high stock volatility than in periods of low stock volatility (*dynamic* flight-to-quality). Across stocks, an overall measure of dynamic flight-to-quality is then

$$\mathsf{FQ}_2 \doteq \sum_{i=1,\ldots,N} \mathbf{1}\,(t_i > 2.33) - \sum_{i=1,\ldots,N} \mathbf{1}\,(t_i < -2.33)\,,$$

with t-statistics defined as $t_i \doteq \hat{c}_3^{(i)} / \mathsf{HAC}[\hat{c}_3^{(i)}]$, where $\hat{c}_3^{(i)}$ is the OLS estimator of $c_3^{(i)}$. Positive values of FQ_2 indicate dynamic flight-*to*-quality, whereas negative values are associated with dynamic flight-*from*-quality.

In line with the prediction of flight-*to*-quality, the observed FQ_2 is always positive (see Table 2). Moreover, it is always significant against the null hypothesis of MGARCH-VAR dynamics. We notice that the misspecified MGARCH-VAR model still captures unexpected flight-*from*-quality, as the bootstrap mean of FQ_2 is negative and significant. However, differently from the test based on FQ_1, the test based on FQ_2 has power against the misspecification as the MGARCH-VAR model is always rejected.

8 Conclusions

Market and funding illiquidity are two essential features of financial markets, as they are related to the inner functioning of the trading activity. Providing a clear definition and estimation of them is challenging, as they are elusive and multifaceted. In this paper, we give notions of daily realized market and funding (ill-)liquidity built from a micro-founded structural model of price formation. Our modeling approach aligns with previous contributions and extends across different directions, allowing uninformed and partially informed traders, the latter facing idiosyncratic and systematic transaction costs. Suitable parametric assumptions allow the likelihood of the model to be written explicitly and, accordingly, to derive efficient estimators of the model parameters. Once estimated with high-frequency intra-day transaction prices (the unique kind of data necessary to carry out the estimation), the model provides daily estimates of stocks' volatility, market, and funding illiquidities. Using a large dataset of NYSE-listed stocks, we thus obtain estimated daily time series of stocks' volatilities, market illiquidities, and funding illiquidity, the latter being assumed to be systemic, i.e., common across all stocks. We study the reciprocal influence among these estimated time series by testing four conjectures, summarized from previous theoretical literature results. In this respect, our empirical exercise delivers robust statistical evidence in favor of four stylized facts: 1) stocks' volatilities and illiquidities (market and funding) show positive co-movements, 2) these co-movements are asymmetric in the sense that they are stronger when funding liquidity dries-up, 3) market illiquidity is higher for higher volatility stocks and 4) when funding liquidity dries-up, high-volatility stocks are more sensitive to changes in funding liquidity than low-volatility stocks.

References

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets 5*(1), 31–56.

Amihud, Y. and H. Mendelson (1987). Trading mechanism and stock retuns: An empirical investigation. *The Journal of Finance 42*(3), 533–553.

Andersen, T. G. and T. Teräsvirta (2009). *Realized Volatility*, pp. 555–575. Springer Berlin Heidelberg.

Bandi, F., A. Kolokolov, D. Pirino, and R. Renò (2020). Zeros. *Management Science 66*(8), 3466–3479.

Bandi, F. M., D. Pirino, and R. Renó (2017). Excess Idle Time. *Econometrica 85*(6), 1793–1846.

Bandi, F. M., D. Pirino, and R. Renó (2024). Systematic staleness. *Journal of Econometrics 238*, 105522.

Bauwens, L., L. Grigoryeva, and J.-P. Ortega (2016). Estimation and empirical performance of non-scalar dynamic conditional correlation models. *Computational Statistics & Data Analysis 100*, 17–36.

Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006). Multivariate garch models: A survey. *Journal of Applied Econometrics 21*(1), 79–109.

Biais, B., L. Glosten, and C. Spatt (2005). Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets 8*(2), 217–264.

Billio, M., M. Caporin, and M. Gobbo (2006, March). Flexible Dynamic Conditional Correlation multivariate GARCH models for asset allocation. *Applied Financial Economics Letters 2*(2), 123–130.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy 96*(1), 116–131.

Boudt, K., E. C. Paulus, and D. W. Rosenthal (2017). Funding liquidity, market liquidity and ted spread: A two-regime model. *Journal of Empirical Finance 43*, 143–158.

Boussama, F., F. Fuchs, and R. Stelzer (2011). Stationarity and geometric ergodicity of bekk multivariate garch models. *Stochastic Processes and their Applications 121*(10), 2331–2360.

Brunnermeier, M. K. (2009, March). Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic Perspectives 23*(1), 77–100.

Brunnermeier, M. K. and L. H. Pedersen (2009). Market liquidity and funding liquidity. *Review of Financial studies 22*(6), 2201–2238.

Chordia, T., R. Roll, and A. Subrahmanyam (2000). Commonality in liquidity. *Journal of Financial Economics 56*(1), 3–28.

Comerton-Forde, C., T. Hendershott, C. M. Jones, P. C. Moulton, and M. S. Seasholes (2010). Time variation in liquidity: The role of market-maker inventories and revenues. *The Journal of Finance 65*(1), 295–331.

Ding, Z. and R. F. Engle (2001). Large scale conditional covariance matrix modeling. *Academia Economic Papers 29*(1), 157–184.

Easley, D. and M. O'Hara (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics 19*(1), 69–90.

Efron, B. and R. J. Tibshirani (1993). *An Introduction to the Bootstrap*. Number 57 in Monographs on Statistics and Applied Probability. Boca Raton, Florida, USA: Chapman & Hall/CRC.

Engle, R. and J. Mezrich (1996). Garch for groups. *Risk 9*, 36–40.

Farouh, M. and R. Garcia (2021). Funding conditions, transaction costs and the performance of anomalies. *SSRN*.

Florackis, C., A. Gregoriou, and A. Kostakis (2011). Trading frequency and asset pricing on the london stock exchange: Evidence from a new price impact ratio. *Journal of Banking & Finance 35*(12), 3335–3350.

Fong, K. Y., C. W. Holden, and C. Trzcinka (2014). What are the best liquidity proxies for global research? Working Paper.

Fontaine, J.-S. and R. Garcia (2012). Bond liquidity premia. *The Review of Financial Studies 25*(4), 1207–1254.

Frazzini, A. and L. H. Pedersen (2014). Betting against beta. *Journal of Financial Economics 111*(1), 1–25.

Gârleanu, N. and L. H. Pedersen (2011, 04). Margin-based asset pricing and deviations from the law of one price. *The Review of Financial Studies 24*(6), 1980–2022.

Geanakoplos, J. (2003). Liquidity, default, and crashes: Endogenous contracts in general equilibrium. In M. Dewatripont, L. P. Hansen, and S. J. Turnovsky (Eds.), *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, Volume 2 of *Econometric Society Monographs*, pp. 170–205. Cambridge University Press.

Glosten, L. and P. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*(1), 71–100.

Golez, B., J. C. Jackwerth, and A. Slavutskaya (2017). Funding illiquidity implied by s&p 500 derivatives. *Econometric Modeling: Derivatives eJournal*.

Goyenko, R., C. Holden, and C. Trzcinka (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics 92*(2), 153–181.

Gromb, D. and D. Vayanos (2002). Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics 66*(2), 361–407. Limits on Arbitrage.

Gromb, D. and D. Vayanos (2010). A model of financial market liquidity based on intermediary capital. *Journal of the European Economic Association 8*(2-3), 456–466.

Hameed, A., W. Kang, and S. Viswanathan (2010). Stock market declines and liquidity. *The Journal of Finance 65*(1), 257–293.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Hasbrouck, J. (2009). Trading costs and returns for us equities: Estimating effective costs from daily data. *The Journal of Finance 64*(3), 1445–1477.

Hasbrouck, J. and T. Ho (1987). Order arrival, quote behavior, and the return-generating process. *The Journal of Finance 42*(4), 1035–1048.

Hu, G. X., J. Pan, and J. Wang (2013). Noise as information for illiquidity. *The Journal of Finance 68*(6), 2341–2382.

Huang, B., Y. Huan, L. D. Xu, L. Zheng, and Z. Zou (2019). Automated trading systems statistical and machine learning methods and hardware implementation: a survey. *Enterprise Information Systems 13*(1), 132–144.

Karolyi, G. A., K.-H. Lee, and M. A. van Dijk (2012). Understanding commonality in liquidity around the world. *Journal of Financial Economics 105*(1), 82–112.

Kitagawa, G. (1987). Non–gaussian state-—space modeling of nonstationary time series. *Journal of the American Statistical Association 82*(400), 1032–1041.

Kotz, S. and S. Nadarajah (2004, December). *Multivariate T-Distributions and Their Applications*. Number 9780521826549 in Cambridge Books. Cambridge University Press.

Kyle, P. (1985). Continuous auctions and insider trading. *Econometrica 43*, 1315–1335.

Le, H. and A. Gregoriou (2020). How do you capture liquidity? a review of the literature on low-frequency stock liquidity. *Journal of Economic Surveys 34*(5), 1170–1186.

Lesmond, D., J. Ogden, and C. Trzcinka (1999). A new estimate of transaction costs. *Review of Financial Studies 12*(5), 1113.

Liu, W. (2006). A liquidity-augmented capital asset pricing model. *Journal of Financial Economics 82*(3), 631–671.

Macchiavelli, M. and X. A. Zhou (2022). Funding liquidity and market liquidity: The broker-dealer perspective. *Management Science 68*(5), 3379–3398.

MacKinnon, J. G. (2009). *Bootstrap Hypothesis Testing*, Chapter 6, pp. 183–213. John Wiley & Sons, Ltd.

Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets 3*(3), 205–258.

Mancini, L., A. Ranaldo, and J. Wrampelmeyer (2013). Liquidity in the foreign exchange market: Measurement, commonality, and risk premiums. *The Journal of Finance 68*(5), 1805–1841.

Mancini-Griffoli, T. and A. Ranaldo (2011). Limits to arbitrage during the crisis: Funding liquidity constraints and covered interest parity. *Working Paper, at SSRN: https://ssrn.com/abstract=1549668*.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Pakel, C., N. Shephard, K. Sheppard, and R. F. Engle (2021). Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics 39*(3), 652–668.

Park, Y.-H. (2015). Price dislocation and price discovery in the SP 500 options and vix derivatives markets. *SSRN*.

Qian, X., L. H. Tam, and B. Zhang (2014). Systematic liquidity and the funding liquidity hypothesis. *Journal of Banking & Finance 45*, 304–320.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance 39*(4), 1127–1139.

Silvennoinen, A. and T. Teräsvirta (2009). *Multivariate GARCH Models*, pp. 201–229. Springer Berlin Heidelberg.

A Proofs

Proof of Theorem 3.1. For any $t = 1, \dots, T$, the conditional expectations of g_t satisfies: i) $\mathsf{E}\{g_t \mid g_{t-1} = -g\} = -g\,\mathsf{T}_{11} + 0\,\mathsf{T}_{12} + g\,\mathsf{T}_{13} = -g\omega + 0(1-\omega) + 0g = -\omega g$; ii) $\mathsf{E}\{g_t|g_{t-1} = 0\} = -g\,\mathsf{T}_{21} + 0\,\mathsf{T}_{22} + g\,\mathsf{T}_{23} = -g(1-\psi)/2 + 0\psi + g(1-\psi)/2 = 0 = \omega 0$, and iii) $\mathsf{E}\{g_t|g_{t-1} = g\} = -g\,\mathsf{T}_{31} + 0\,\mathsf{T}_{32} + g\,\mathsf{T}_{33} = -g0 + 0(1-\omega) + \omega g = \omega g$. Hence, in general, we have $\mathsf{E}\{g_t \mid g_{t-1}\} = \omega g_{t-1}$, which means that the process $\{u_t\}$, with $u_t \doteq g_t - \omega g_{t-1}$, is a martingale difference. Since $\{v_t\}$ is a Gaussian white noise independent of $\{g_t\}$, the processes $\{(u_t, v_t)\}$ and $\{(\eta_t, \zeta_t)\}$ are martingale differences, which proves that $\mathsf{E}\left\{\eta_t|\bar{\mathcal{E}}_{t-1}\right\} = \mathsf{E}\left\{\zeta_t|\bar{\mathcal{E}}_{t-1}\right\} = 0$. Then, to prove that assumptions (3) and (5-6) hold, with $(\sigma^2_\eta, \sigma^2_\zeta) = (a, b)$, it remains to prove that

$$\mathsf{E}\left\{\eta_t^2\right\} = a, \qquad \mathsf{E}\left\{\zeta_t^2\right\} = b, \qquad \text{and} \qquad \mathsf{E}\{\eta_t\zeta_t\} = 0. \tag{30}$$

Given the transition matrix of $\{g_t\}$, and the distribution of g_0, the process $\{g_t\}$ is strictly stationary with stationary low as in (14) (See, e.g., Hamilton, 1994, Chapter 22.2). [18] The stationary low is symmetric around zero, and, hence, $\mathsf{E}\{g_t\} = 0$. Therefore, applying (12) and (14), the variance of $\{g_t\}$ coincides with the second moment of $\{g_t\}$,

$$\mathsf{E}\left\{g_t^2\right\} = g^2\mathsf{P}\{g_0 = -g\} + 0\mathsf{P}\{g_0 = 0\} + g^2\mathsf{P}\{g_0 = g\} = \omega^2(a+b)/(1-\omega^2).$$

On the other hand, since $g_t = \omega g_{t-1} + u_t$ is a stationary AR(1) process, it holds that $\mathsf{E}\left\{g_t^2\right\} = \mathsf{E}\left\{u_t^2\right\}/(1-\omega^2)$ (Hamilton, 1994). It follows that $\mathsf{E}\left\{u_t^2\right\} = \omega^2(a+b)$. Applying $\mathsf{E}\left\{v_t^2\right\} \doteq (ab)/(a+b)$, the independence of $\{v_t\}$ and $\{u_t\}$, and the definition of $\{(\eta_t, \zeta_t)\}$ in (10), simple calculations yields (30), which completes the proof of the first part of the theorem. As for the second part, recalling that $\bar{e}_t = \bar{e}_{t-1} + \eta_t$, and noticing that $m_t = m_{t-1} + \delta(\bar{e}_t - m_{t-1}) + (1-\delta)\zeta_t = \delta\bar{e}_{t-1} + \delta\eta_t + \omega m_{t-1} + \omega\zeta_t$, we can write $\bar{e}_t - m_t = \omega(\bar{e}_{t-1} - m_{t-1}) + \omega(\eta_t - \zeta_t)$, where $\omega(\eta_t - \zeta_t) = \omega(\gamma_\eta u_t + \gamma_\zeta u_t) = \omega(\gamma_\eta + \gamma_\zeta)u_t = \omega(1/\omega)u_t = u_t$. Hence, $\bar{e}_t - m_t = \omega(\bar{e}_{t-1} - m_{t-1}) + u_t$. Since $g_t = \omega g_{t-1} + u_t$, we have $\bar{e}_t - m_t = g_t$. □

Proof of Theorem 3.2. From the definition of $\{s_t\}$ given in Section 3.2, for $t > 0$ we have, a.s.,

$$s_t = 0 \Longleftrightarrow p_t = p_{t-1}, \quad s_t = -s \Longleftrightarrow m_t = p_t + s \quad \text{and} \quad s_t = s \Longleftrightarrow m_t = p_t - s. \tag{31}$$

Hence, letting $f^{(m)}_{t-1}(m_t)$ denote the probability density function (PDF) of $m_t|\mathcal{F}_{t-1}$, we can write $f_{t-1}(p_t)$ as

$$\begin{aligned} f_{t-1}(p_t) \quad = \quad & \mathsf{P}\{s_t = \;\; 0 \mid \mathcal{F}_{t-1}\} \cdot \mathbf{1}(p_t = p_{t-1}) + \mathsf{P}\{s_t = -s \mid \mathcal{F}_{t-1}\} \cdot f^{(m)}_{t-1}(p_t + s) \\ & + \mathsf{P}\{s_t = s \mid \mathcal{F}_{t-1}\} \cdot f^{(m)}_{t-1}(p_t - s). \end{aligned}$$

[18] The vector of the stationary probabilities coincides with the left eigenvector of T associated with the unit eigenvalue.

Recall that whenever a price change occurs, we have $p_t = m_t + s_t$, with $s_t = \pm s$. Whence, applying the law of total probability, for $q = 0, 1, \dots$ and for $t_q < t \leq t_{q+1}$ we can write

$$\begin{aligned} f_{t-1}(p_t) \quad = \quad & \sum_{s_t=0,\,(g_t,s_{t_q},g_{t_q})\in G\times S\times G} P_{t-1}(s_t, g_t, s_{t_q}, g_{t_q}) \cdot \mathbf{1}(p_t = p_{t-1}) + \\ & \sum_{s_t=-s,\,(g_t,s_{t_q},g_{t_q})\in G\times S\times G} P_{t-1}(s_t, g_t, s_{t_q}, g_{t_q}) \cdot f^{(m)}_{t-1}(p_t + s|g_t, s_{t_q}, g_{t_q}) + \\ & \sum_{s_t=+s,\,(g_t,s_{t_q},g_{t_q})\in G\times S\times G} P_{t-1}(s_t, g_t, s_{t_q}, g_{t_q}) \cdot f^{(m)}_{t-1}(p_t - s|g_t, s_{t_q}, g_{t_q}), \end{aligned} \tag{32}$$

where $f^{(m)}_{t-1}(m_t|g_t, s_{t_q}, g_{t_q})$ denotes the PDF of $m_t|(g_t, s_{t_q}, g_{t_q}, \mathcal{F}_{t-1})$. Suppose, for a moment, that the latter PDF satisfies

$$f^{(m)}_{t-1}(m_t|g_t, s_{t_q}, g_{t_q}) = f_G(m_t; \mu^{(m)}_t, \lambda_t), \tag{33}$$

with $\mu^{(m)}_t \doteq p_{t_q} + \rho g_t - s_{t_q} + \xi g_{t_q}$. Rewriting μ_t as $\mu_t = p_{t_q} + s_t + \rho g_t - s_{t_q} + \xi g_{t_q} = \mu^{(m)}_t + s_t$, and applying (31), we could write

$$f^{(m)}_{t-1}(p_t + s|g_t, s_{t_q}, g_{t_q}) = f_G(p_t; \mu_t, \lambda_t), \qquad \text{for} \quad s_t = -s, \tag{34}$$

and

$$f^{(m)}_{t-1}(p_t - s|g_t, s_{t_q}, g_{t_q}) = f_G(p_t; \mu_t, \lambda_t), \qquad \text{for} \quad s_t = s. \tag{35}$$

Then, replacing $f^{(m)}_{t-1}(p_t + s|g_t, s_{t_q}, g_{t_q})$ and $f^{(m)}_{t-1}(p_t - s|g_t, s_{t_q}, g_{t_q})$ in (32) with $f_G(p_t; \mu_t, \lambda_t)$, would prove the theorem. Hence, all we need to prove the theorem is that equation (33) holds. This is done as follows.

Recalling that $m_t = m_{t-1} + \delta(\bar{e}_t - m_{t-1}) + (1 - \delta)\zeta_t$, where $\delta \in (0, 1)$, we write $m_t = m_{t-1} + \delta(\bar{e}_t - m_t + m_t - m_{t-1}) + (1 - \delta)\zeta_t$, that we rearrange as $(1-\delta)m_t = (1-\delta)m_{t-1} + \delta(\bar{e}_t - m_t) + (1-\delta)\zeta_t$. Dividing both sides by $1 - \delta$, and using the three equalities 1) $\omega = 1 - \delta$, 2) $\zeta_t = \nu_t - \gamma_\zeta u_t$ and 3) $u_t = g_t - \omega g_{t-1}$, we can write $m_t = m_{t-1} + (\delta/\omega)g_t + \nu_t - \gamma_\zeta(g_t - \omega g_{t-1})$, or, equivalently, $m_t = m_{t-1} + (\delta/\omega - \gamma_\zeta)g_t + (\gamma_\zeta\omega)g_{t-1} + \nu_t$. Having defined $\rho = \delta/\omega - \gamma_\zeta$ and $\xi = \gamma_\zeta\omega$, we can write m_t as $m_t = m_{t-1} + \rho g_t + \xi g_{t-1} + \nu_t$. Then, by backward substitutions, we can write m_t as

$$m_t = m_{t_q} + \rho g_t + \xi g_{t_q} + (\rho + \xi) \sum_{j=t_q+1}^{t-1} g_j + \sum_{j=t_q+1}^{t} \nu_j.$$

Recalling that t is such that $t_q < t \leq t_{q+1}$, from time $t_q + 1$ to time $t - 1$ the transaction price is constant, that is,

$$p_{t_q+1} = p_{t_q+2} = \cdots = p_{t-2} = p_{t-1}.$$

This implies in turns that $g_{t_q+1} = g_{t_q+2} = \cdots = g_{t-2} = g_{t-1} = 0$, which yields $(\rho+\xi)\sum_{j=t_q+1}^{t-1} g_j = 0$. Hence, we can write m_t as

$$m_t = m_{t_q} + \rho g_t + \xi g_{t_q} + \sum_{j=t_q+1}^{t} \nu_j.$$

Recalling that $p_{t_q} = m_{t_q} + s_{t_q}$, we can further write

$$m_t = p_{t_q} - s_{t_q} + \rho g_t + \xi g_{t_q} + \sum_{j=t_q+1}^{t} \nu_j.$$

Now, conditionally on $\mathcal{F}_{t_q}$, the distribution of $\sum_{j=t_q+1}^{t} \nu_j$ is $N(0, \lambda_t)$, with $\lambda_t = (t - t_q)\sigma_\nu^2$. Moreover, $\{\nu_t\}$ is independent of $\{(g_t, s_t)\}$ (see the definitions in Theorem 3.1). Therefore, the distribution of m_t given $(g_t, s_{t_q}, g_{t_q}, \mathcal{F}_{t_q})$ is normal, with mean $p_{t_q} + \rho g_t - s_{t_q} + \xi g_{t_q}$ and variance λ_t, or

$$m_t|(g_t, s_{t_q}, g_{t_q}, \mathcal{F}_{t_q}) \sim N(\mu_t^{(m)}, \lambda_t). \tag{36}$$

Recall that, for $t = t_q + 1, t_q + 2, \ldots, t_{q+1} - 1$, the price sequence is stale, i.e., it remains constant. Accordingly, we can replace $\mathcal{F}_{t_q}$ in (36) with $\mathcal{F}_{t-1}$, getting

$$m_t|(g_t, s_{t_q}, g_{t_q}, \mathcal{F}_{t-1}) \sim N(\mu_t^{(m)}, \lambda_t),$$

which proves (33). □

B Forward recursion for the mixture weights

The mixture weights required for the computation of $f_{t-1}(p_t)$ can be computed in closed form via Bayesian forward recursion, as follows.

For $q = 0, 1 \ldots$ and for $t_q < t \leq t_{q+1}$, we define the process $\{a_t\}$ setting $a_t \doteq (s_t, g_t, s_{t_q}, g_{t_q})$, with the convention that $a_0 \doteq (s_0, g_0, s_{-1}, g_{-1}) \doteq (0, 0, 0, 0)$.[19] By means of this notation, $f_{t-1}(p_t)$ can be written, more compactly, as

$$f_{t-1}(p_t) = \sum_{a_t \in S\times G\times S\times G} P_{t-1}(a_t) \cdot f_{t-1}(p_t|a_t),$$

[19] For instance, assuming that, for a given q, we have $t_{q-1} = 3$, $t_q = 6$, and $t_{q+1} = 8$, we get, for $t_{q-1} < t \leq t_q$ and $t_q < t \leq t_{q+1}$,

$$\begin{aligned}
t = 4 &\Longrightarrow a_4 = (s_4, g_4, s_{t_{q-1}}, g_{t_{q-1}}) = (s_4, g_4, s_3, g_3),\\
t = 5 &\Longrightarrow a_5 = (s_5, g_5, s_{t_{q-1}}, g_{t_{q-1}}) = (s_5, g_5, s_3, g_3),\\
t = 6 &\Longrightarrow a_6 = (s_{t_q}, g_{t_q}, s_{t_{q-1}}, g_{t_{q-1}}) = (s_6, g_6, s_3, g_3),\\
t = 7 &\Longrightarrow a_7 = (s_7, g_7, s_{t_q}, g_{t_q}) = (s_7, g_7, s_6, g_6),\\
t = 8 &\Longrightarrow a_8 = (s_8, g_8, s_{t_q}, g_{t_q}) = (s_8, g_8, s_6, g_6).
\end{aligned}$$

where $f_{t-1}(p_t|a_t) \doteq \mathbf{1}(p_t = p_{t-1})$ for $s_t = 0$, and $f_{t-1}(p_t|a_t) \doteq \mathbf{1}(p_t \neq p_{t-1}) \cdot f_G(p_t; \mu_t, \lambda_t)$ for $s_t = \pm s$. The mixture weights are now denoted with $P_{t-1}(a_t)$, for $a_t \in S \times G \times S \times G$.

Assuming that the posterior distribution, $P_{t-1}(a_{t-1}) \doteq P(a_{t-1}|\mathcal{F}_{t-1})$, is known, and that the transition distribution, $P_{t-1}(a_t|a_{t-1}) \doteq P(a_t|a_{t-1}, \mathcal{F}_{t-1})$ is known, the mixture weights are computed via prediction step, as

$$P_{t-1}(a_t) = \sum_{a_{t-1} \in A} P_{t-1}(a_t|a_{t-1}) \cdot P_{t-1}(a_{t-1}).$$

The new posterior distribution is computed next via the filtering step, as

$$P_t(a_t) = \frac{f_{t-1}(p_t|a_t) \cdot P_{t-1}(a_t)}{\sum_{\tilde{a}_t \in A} f_{t-1}(p_t|\tilde{a}_t) \cdot P_{t-1}(\tilde{a}_t)}. \tag{37}$$

Since a_0 is known, the initial posterior distribution, $P_0(a_0)$, is known. Therefore, all we need to run the entire forward recursion in closed form is a closed-form expression of the transition distribution, $P_{t-1}(a_t|a_{t-1})$, for all $t \geq 1$. This closed-form expression is derived in two steps. First we derive a closed-form expression for $P_{t-1}(s_t, g_t|s_{t-1}, g_{t-1}) \doteq P(s_t, g_t|s_{t-1}, g_{t-1}, \mathcal{F}_{t-1})$. Then, we write $P_{t-1}(a_t|a_{t-1})$ in terms of the closed form of $P_{t-1}(s_t, g_t|s_{t-1}, g_{t-1})$.

As for the closed form of $P_{t-1}(s_t, g_t|s_{t-1}, g_{t-1})$, let us write

$$P_{t-1}(s_t, g_t|s_{t-1}, g_{t-1}) = P_{t-1}(s_t|g_t, s_{t-1}, g_{t-1}) \cdot P_{t-1}(g_t|s_{t-1}, g_{t-1}).$$

The probabilities distributions on the right hand side, $P_{t-1}(s_t|g_t, s_{t-1}, g_{t-1})$ and $P_{t-1}(g_t|s_{t-1}, g_{t-1})$, can be written in closed form as follows.

From assumptions (7–8), it follows that the distribution of s_t given $(g_t, s_{t-1}, g_{t-1}, \mathcal{F}_{t-1})$, turns out to be affected only by g_t. Therefore, $P_{t-1}(s_t|g_t, s_{t-1}, g_{t-1})$ can equivalently be written as $P(s_t|g_t)$. Applying assumptions (7–8), we get $P(s_t|g_t)$ in closed form as in the following scheme,

$$P(s_t|g_t) \quad = \quad \begin{array}{ccc|l} s_t = -s & s_t = 0 & s_t = s & \\ \hline (1+\pi)/2 & 0 & (1-\pi)/2 & g_t = -g \\ (1-\pi)/2 & \pi & (1-\pi)/2 & g_t = 0 \\ (1-\pi)/2 & 0 & (1+\pi)/2 & g_t = g \end{array} \; .$$

Concerning the closed form of $P_{t-1}(g_t|s_{t-1}, g_{t-1})$, we notice that the distribution of g_t given $(s_{t-1}, g_{t-1}, \mathcal{F}_{t-1})$, is unaffected by conditioning on $(s_{t-1}, \mathcal{F}_{t-1})$. Therefore, $P_{t-1}(g_t|s_{t-1}, g_{t-1})$ can equivalently be written as $P(g_t|g_{t-1})$, where $P(g_t|g_{t-1})$ is the transition distribution of g_t given in (13). We can finally write $P_{t-1}(s_t, g_t|s_{t-1}, g_{t-1})$ as

$$P_{t-1}(s_t, g_t|s_{t-1}, g_{t-1}) = P(s_t|g_t) \cdot P(g_t|g_{t-1}) = P(s_t|g_{t-1}), \tag{38}$$

where $P(s_t|g_t)$ and $P(g_t|g_{t-1})$ have the closed form described above. Since $P_{t-1}(s_t, g_t|s_{t-1}, g_{t-1})$ does not depend on t, we shall write $P(s_t, g_t|s_{t-1}, g_{t-1})$ in place of $P_{t-1}(s_t, g_t|s_{t-1}, g_{t-1})$. Notice that $\{(s_t, g_t)\}$ is a Markov process[20]. We will make use of the Markov property of $\{(s_t, g_t)\}$ in a moment.

Having derived the closed-form expression of $P(s_t, g_t|s_{t-1}, g_{t-1})$, we now write $P_{t-1}(a_t|a_{t-1})$ in terms of $P(s_t, g_t|s_{t-1}, g_{t-1})$. In doing this, it is convenient to write $P_{t-1}(a_t|\tilde{a}_{t-1})$ in place of $P_{t-1}(a_t|a_{t-1})$, with $\tilde{a}_{t-1} \doteq (\tilde{s}_{t-1}, \tilde{g}_{t-1}, \tilde{s}_{t_{q-1}}, \tilde{g}_{t_{q-1}}) \in S \times G \times S \times G$. We shall distinguish two cases, depending on either $t = t_q + 1$ or $t_q + 1 < t \leq t_{q+1}$.

If $t = t_q + 1$, we have $a_t|\tilde{a}_{t-1} = (s_{t_q+1}, g_{t_q+1}, s_{t_q}, g_{t_q})\big|(\tilde{s}_{t_q}, \tilde{g}_{t_q}, \tilde{s}_{t_{q-1}}, \tilde{g}_{t_{q-1}})$. For $(s_{t_q}, g_{t_q}) = (\tilde{s}_{t_q}, \tilde{g}_{t_q})$, we can write

$$P_{t-1}(a_t|\tilde{a}_{t-1}) = P_{t_q}(s_{t_q+1}, g_{t_q+1}, s_{t_q}, g_{t_q}|\tilde{s}_{t_q}, \tilde{g}_{t_q}, \tilde{s}_{t_{q-1}}, \tilde{g}_{t_{q-1}}) = P_{t_q}(s_{t_q+1}, g_{t_q+1}|\tilde{s}_{t_q}, \tilde{g}_{t_q}, \tilde{s}_{t_{q-1}}, \tilde{g}_{t_{q-1}}).$$

By the Markov property of $\{(s_t, g_t)\}$, the distribution of $(s_{t_q+1}, g_{t_q+1})|(\tilde{s}_{t_q}, \tilde{g}_{t_q}, \tilde{s}_{t_{q-1}}, \tilde{g}_{t_{q-1}}, \mathcal{F}_{t_q})$ turns out to be unaffected by the conditioning event $(\tilde{s}_{t_{q-1}}, \tilde{g}_{t_{q-1}}, \mathcal{F}_{t_q})$. Therefore, we can write

$$P_{t-1}(a_t|\tilde{a}_{t-1}) = P(s_{t_q+1}, g_{t_q+1}|\tilde{s}_{t_q}, \tilde{g}_{t_q}) = P(s_{t_q+1}|\tilde{g}_{t_q}), \tag{39}$$

where we applied equation (38). For $(s_{t_q}, g_{t_q}) \neq (\tilde{s}_{t_q}, \tilde{g}_{t_q})$, which implies a contradiction, the transition from $\tilde{a}_{t-1}$ to a_t is not possible, yielding $P_{t-1}(a_t|\tilde{a}_{t-1}) = 0$.

If $t_q + 1 < t \leq t_{q+1}$, we have $a_t|\tilde{a}_{t-1} = (s_t, g_t, s_{t_q}, g_{t_q})\big|(\tilde{s}_{t-1}, \tilde{g}_{t-1}, \tilde{s}_{t_q}, \tilde{g}_{t_q})$. For $(s_{t_q}, g_{t_q}) = (\tilde{s}_{t_q}, \tilde{g}_{t_q})$, we can write

$$P_{t-1}(a_t|\tilde{a}_{t-1}) = P_{t-1}(s_t, g_t, s_{t_q}, g_{t_q}|\tilde{s}_{t-1}, \tilde{g}_{t-1}, \tilde{s}_{t_q}, \tilde{g}_{t_q}) = P_{t-1}(s_t, g_t|\tilde{s}_{t-1}, \tilde{g}_{t-1}, \tilde{s}_{t_q}, \tilde{g}_{t_q}).$$

Noticing that $t_q < t - 1$, by the Markov property of $\{(s_t, g_t)\}$ it follows that the distribution of $(s_t, g_t)|(\tilde{s}_{t-1}, \tilde{g}_{t-1}, \tilde{s}_{t_q}, \tilde{g}_{t_q}, \mathcal{F}_{t-1})$ turns out to be unaffected by conditioning on $(\tilde{s}_{t_q}, \tilde{g}_{t_q}, \mathcal{F}_{t-1})$. Therefore, we can write

$$P_{t-1}(a_t|\tilde{a}_{t-1}) = P(s_t, g_t|\tilde{s}_{t-1}, \tilde{g}_{t-1}) = P(s_t|\tilde{g}_{t-1}), \tag{40}$$

where we applied equation (38). For $(s_{t_q}, g_{t_q}) \neq (\tilde{s}_{t_q}, \tilde{g}_{t_q})$, which implies a contradiction, the transition from $\tilde{a}_{t-1}$ to a_t is not possible, yielding $P_{t-1}(a_t|\tilde{a}_{t-1}) = 0$.

In summary, if $(s_{t_q}, g_{t_q}) = (\tilde{s}_{t_q}, \tilde{g}_{t_q})$, we can write $P_{t-1}(a_t|\tilde{a}_{t-1})$ in terms of the closed form of $P(s_t, g_t|s_{t-1}, g_{t-1})$ given in equation (38). This is done by applying either (39) or (40), depending on whether $t = t_q + 1$ or $t_q + 1 < t \leq t_{q+1}$. If $(s_{t_q}, g_{t_q}) \neq (\tilde{s}_{t_q}, \tilde{g}_{t_q})$, we have $P_{t-1}(a_t|\tilde{a}_{t-1}) = 0$.

[20] Some states of $S \times G$ are inaccessible by $\{(s_t, g_t)\}$. In particular, recalling assumptions (7–8), for $t > 0$ we have that $P(s_t, g_t|s_{t-1}, g_{t-1}) = 0$ for all $(s_{t-1}, g_{t-1}) \in S \times G$, if and only if $s_t = 0 \wedge g_t \neq 0$.

C Estimator of $\mathsf{E}\left\{\bar{e}^{(i)}_{d,t} \,\middle|\, \mathcal{F}_T\right\}$

Recall that, for $t = t_{q+1}$, we have $p_t = m_t + s_t$. Since, in general, it holds that $\bar{e}_t = m_t + g_t$, if $t = t_{q+1}$ we can write $\bar{e}_t = m_t + g_t = p_t - s_t + g_t$.[21] Whence, for $t = t_{q+1}$, we have

$$\mathsf{E}\{\bar{e}_t|\mathcal{F}_T\} = \mathsf{E}\{p_t - s_t + g_t|\mathcal{F}_T\} = p_t - \mathsf{E}\{s_t|\mathcal{F}_T\} + \mathsf{E}\{g_t|\mathcal{F}_T\}. \tag{41}$$

The expectations, $\mathsf{E}\{s_t|\mathcal{F}_T\}$ and $\mathsf{E}\{g_t|\mathcal{F}_T\}$, can be computed from the fixed interval smoothing distribution of $\{(s_t, g_t)\}$, denoted with $P_T(s_t, g_t)$. The latter is obtained recursively as

$$P_T(s_t, g_t) = P_t(s_t, g_t) \cdot \sum_{s_{t+1}, g_{t+1}} \frac{P(s_{t+1}, g_{t+1}|s_t, g_t) \cdot P_T(s_{t+1}, g_{t+1})}{\sum_{\tilde{s}_t, \tilde{g}_t} P(s_{t+1}, g_{t+1}|\tilde{s}_t, \tilde{g}_t) \cdot P_t(\tilde{s}_t, \tilde{g}_t)},$$

$t = T-1, t-2, \ldots, 1$, where $P(s_{t+1}, g_{t+1}|s_t, g_t)$ and $P_t(s_t, g_t)$ are the transition and filtering distributions of $\{(s_t, g_t)\}$ (Kitagawa, 1987). The closed form of $P(s_{t+1}, g_{t+1}|s_t, g_t)$ is given in equation (38) of Appendix B. The closed form of $P_t(s_t, g_t)$ can be derived by marginalization from the closed form of the filtering distribution of $\{a_t\}$, also derived in Appendix B.

For $q = 0, 1, \ldots$ and for $t_q < t \leq t_{q+1}$, we approximate the smoothing mean, $\mathsf{E}\{\bar{e}_t|\mathcal{F}_T\}$, with the linear interpolant

$$\hat{e}_t \doteq \mathsf{E}\left\{\bar{e}_{t_q}|\mathcal{F}_T\right\} + \frac{\mathsf{E}\left\{\bar{e}_{t_{q+1}}|\mathcal{F}_T\right\} - \mathsf{E}\left\{\bar{e}_{t_q}|\mathcal{F}_T\right\}}{t_{q+1} - t_q}(t - t_q),$$

where $\mathsf{E}\left\{\bar{e}_{t_q}|\mathcal{F}_T\right\}$ and $\mathsf{E}\left\{\bar{e}_{t_{q+1}}|\mathcal{F}_T\right\}$ are the exact smoothing mean computed applying (41). If $t = t_{q+1}$,[22] the approximation is exact, or $\hat{e}_t = \mathsf{E}\{\bar{e}_t|\mathcal{F}_T\}$.

D The estimation of the MGARCH-VAR model

The bivariate composite likelihood adopted here is defined as

$$\mathcal{L}(L, C, W, A, B) \doteq \sum_{i<j=2,\ldots,M} \mathcal{L}^{(i,j)}\left(L, C, W_{[i,j]}, A_{[i,j]}, B_{[i,j]}\right),$$

where

$$\mathcal{L}^{(i,j)}\left(L, C, W_{[i,j]}, A_{[i,j]}, B_{[i,j]}\right) \doteq \sum_{d=1}^{D} -\frac{1}{2}\left(4\pi + \log\left|H_d^{[i,j]}\right| + U_d^{[i,j]\top}\left(H_d^{[i,j]}\right)^{-1} U_d^{[i,j]}\right)$$

[21] To avoid clutter, we drop the stock index, (i), and the day index, d, from the notation.

[22] That is, whenever a price change occurs.

is the bivariate Gaussian quasi-log-likelihood associated with the bivariate time series $\left\{U_d^{[i,j]} \mid d = 1, \ldots, D\right\}$, with $U_d^{[i,j]} \doteq (U_d^{(i)}, U_d^{(j)})^\top$. Applying (22), the 2×2 conditional covariance matrix $H_d^{[i,j]}$ satisfies

$$H_d^{[i,j]} = W_{[i,j]} \odot (\mathbb{I}_{2,2} - A_{[i,j]} - B_{[i,j]}) + A_{[i,j]} \odot U_{d-1}^{[i,j]} U_{d-1}^{[i,j]\top} + B_{[i,j]} \odot H_{d-1}^{[i,j]},$$

where $W_{[i,j]}$, $A_{[i,j]}$, and $B_{[i,j]}$, denote the appropriate submatrices of W, A, and B.[23]

After replacing Y_d with $\hat{Y}_d$, we fit the MGARCH-VAR parameters in 5 steps.

1. Compute the optimal value, $\hat{L}$, via minimum-AIC under the convenience assumption of iid Gaussian innovations[24].

2. Compute $\hat{C} \doteq \{\hat{c}_0, \hat{C}_1, \ldots, \hat{C}_{\hat{L}}\}$ via OLS.

3. Compute $\hat{W}$ as the sample second moment of the OLS residuals, $\hat{W} \doteq D^{-1} \sum_{d=1}^{D} \hat{U}_d \hat{U}_d^\top$.[25]

4. Compute $\left\{\hat{A}, \hat{B}\right\}$ as a maximizer of the composite likelihood,

$$\left\{\hat{A}, \hat{B}\right\} \doteq \operatorname{argmax}_{\{A,B\}} \mathcal{L}\left(\hat{L}, \hat{C}, \hat{W}, A, B\right),$$

   subject to positive definiteness of A, B and $\hat{W} \odot (\mathbb{I}_{M,M} - A - B)$.[26]

5. Compute $\hat{\kappa}$ as the ML estimator of κ for fixed $(L, C, W, A, B) = (\hat{L}, \hat{C}, \hat{W}, \hat{A}, \hat{B})$. This means to set

$$\hat{\kappa} \doteq \operatorname{argmax}_{\kappa > 2} \sum_{d=1}^{D} \left\{ \log \frac{\Gamma[(M+\kappa)/2]}{\Gamma[\kappa/2](\kappa\pi)^{M/2}} - \frac{1}{2} \log |\hat{S}_d| - \frac{M+\kappa}{2} \log\left(1 + \frac{1}{\kappa} \hat{U}_d^\top \hat{S}_d^{-1} \hat{U}_d\right)\right\},$$

   with $\hat{S}_d \doteq \hat{H}_d(\kappa - 2)/\kappa$ and $\hat{H}_d \doteq \hat{W} \odot (\mathbb{I}_{M,M} - \hat{A} - \hat{B}) + \hat{A} \odot [\hat{U}_{d-1} \hat{U}_{d-1}^\top] + \hat{B} \odot \hat{H}_{d-1}$.

Under standard regularity conditions, each step is consistent conditionally on the estimation output of the previous steps. The OLS estimator of the VAR coefficients computed at step 2 is consistent even if the optimal lag is overestimated.

[23] The dependence of $\mathcal{L}^{(i,j)}$ on (L, C) arises through $U_d = Y_d - c_0 - C_1 Y_{d-1} - C_2 Y_{d-2} - \cdots - C_L Y_{d-L}$.

[24] This entails to minimize, over $L = 0, 1, \ldots$, the objective function $\log |D^{-1} \sum_{d=1}^{D} \hat{U}_d \hat{U}_d^\top| + (M + LM^2)(2/D)$, where $\left\{\hat{U}_d\right\}$ is the time series of the OLS residuals from the VAR(L) model, and $(M + LM^2)$ is the number of parameters in the VAR model.

[25] This step is the so-called *variance targeting* introduced by Engle and Mezrich (1996).

[26] The required positive definite constraint is not of immediate implementation in the empirical calculations. Here, it is imposed via reparameterizations and penalizations of the objective function. A sophisticated approach, based on a Bregman-proximal trust-region method, is proposed in Bauwens et al. (2016).

RECENT PUBLICATIONS BY *CEIS Tor Vergata*

On The Nonlinearity of the Finance and Growth Relation: the Role of Human Capital
Alberto Bucci, Boubacar Diallo and Simone Marsiglio
CEIS *Research Paper*, 567 November 2023

Distance Work and Life Satisfaction after the COVID-19 Pandemics
Leonardo Becchetti, Gianluigi Conzo and Fabio Pisani
CEIS *Research Paper*, 566 September 2023

Scitovsky Was Right…and There Is More: Comfort Goods, Stimulus Goods, Education and Subjective Wellbeing
Leonardo Becchetti and Chiara Lubicz
CEIS Research Paper, 565 August 2023

Information Campaigns and migration perceptions
Erminia Florio
CEIS Research Paper, 564 August 2023

Is Self-Employment for Migrants? Evidence from Italy
Marianna Brunetti and Anzelika Zaiceva
CEIS Research Paper, 563 July 2023

Italy in the Great Divergence: What Can We Learn from Engel's Law?
David Chilosi and Carlo Ciccarelli
CEIS Research Paper, 562 July 2023

Are Anti-Federalism and Republicanism the Way Forward for a United States of Europe? Lessons from American History
Dandan Hong, Lorenzo Pecchi and Gustavo Piga
CEIS Research Paper, 561 July 2023

How Much Is Too Much? A Methodological Investigation Of The Literature On Alcohol Consumption
Stefano Castriota, Paolo Frumento and Francesco Suppressa
CEIS Research Paper, 560 July 2023

Band-Pass Filtering with High-Dimensional Time Series
Alessandro Giovannelli, Marco Lippi and Tommaso Proietti
CEIS Research Paper, 559 June 2023

Deflation by Expenditure Components: A Harmless Adjustment?
Nicola Amendola, Giulia Mancini, Silvia Redaelli and Giovanni Vecchi
CEIS Research Paper, 558 June 2023

DISTRIBUTION

Our publications are available online at www.ceistorvergata.it

DISCLAIMER

The opinions expressed in these publications are the authors' alone and therefore do not necessarily reflect the opinions of the supporters, staff, or boards of CEIS Tor Vergata.

MEDIA INQUIRIES AND INFORMATION

For media inquiries, please contact Barbara Piazzi at +39 06 72595652/01 or by e-mail at piazzi@ceis.uniroma2.it. Our web site, www.ceistorvergata.it, contains more information about Center's events, publications, and staff.

DEVELOPMENT AND SUPPORT

For information about contributing to CEIS Tor Vergata, please contact at +39 06 72595601 or by e-mail at segr.ceis@economia.uniroma2.it