Testing liquidity: A statistical theory based on asset staleness

Published Version:
Pirino Davide, Pollastri Alessandro, Trapin Luca (In stampa/Attività in corso). Testing liquidity: A statistical theory based on asset staleness. ECONOMETRICS AND STATISTICS, -, 1-18 [10.1016/j.ecosta.2022.07.002].

Availability:
This version is available at: https://hdl.handle.net/11585/914687 since: 2023-02-10

Published:
DOI: http://doi.org/10.1016/j.ecosta.2022.07.002

Testing liquidity: a statistical theory based on asset staleness

Davide Pirino[a], Alessandro Pollastri[b], Luca Trapin[c,*]

[a] *Università degli Studi di Roma "Tor Vergata", Dipartimento di Economia e Finanza, Via Columbia 2, 00173, Roma and Scuola Normale Superiore, Piazza dei Cavalieri 7, 56123, Pisa, Italy. E-mail: davide.pirino@gmail.com*

[b] *De Nederlandsche Bank, Spaklerweg 4, 1096BA, Amsterdam, The Netherlands. Email: A.Pollastri@dnb.nl*

[c] *Department of Statistical Sciences, University of Bologna, Via delle Belle Arti 41, 40126 Bologna, Italy. E-mail: luca.trapin@unibo.it*

Abstract

Using asset staleness as liquidity proxy, two novel test statistics that allow to make inference on the level of liquidity of an asset and on the difference in liquidity between two assets are proposed. The (in-fill) asymptotic properties of the tests are established, and correct procedures to use the tests in multiple testing are provided. A simulation study confirms that the newly defined tests show desirable finite sample properties. Two applications show how the tests can be used for the investor's asset allocation problem in a high-dimensional setting.

Keywords: Asset allocation; Hypothesis testing; Liquidity; Staleness.

1. Introduction

Market liquidity is a central aspect of modern finance, and the financial economics literature has made considerable efforts toward defining and measuring liquidity. The statistical uncertainty surrounding liquidity measures naturally raises issues that are relevant for investments, trading decisions and policy implementation. Consider an investment manager who wants to allocate funds toward equally (il)liquid assets in order to build a portfolio with the desired exposure to illiquidity and benefit from the illiquidity premium. This portfolio strategy requires to know the liquidity level of an asset. Similarly, consider a

*Corresponding author.

policymaker who wants to assess whether a market reform has an impact on the liquidity of an asset. This requires to know whether the liquidity of the asset has changed after the regulatory intervention. As we can only measure liquidity with uncertainty, statistical inference is necessary to answer these questions.

We provide a formal testing procedure, exclusively based on observed transaction prices, to infer the level of (a proxy of) assets' liquidity and to test whether two assets present the same level of liquidity. We build on the econometric framework for *stale* prices of Bandi et al. (2017). Under the null of an Itô-semimartingale they derive the asymptotic distribution of the percentage of log-price variations smaller than an asymptotically vanishing threshold, defining a new economic indicator named *idle time*. Empirically, Bandi et al. (2017) found that the Itô-semimartingale null is conclusively rejected in favor of a frictional alternative in which prices update less frequently than what is assumed by standard asset pricing models in continuous time. In Bandi et al. (2020b), the same authors show that this market feature has a systematic component, i.e. temporary market freezing may occur simultaneously across many assets.

Classical models of price formation provide an economic rationale for the existence of stale prices, suggesting that in the presence of trading costs, rational investors with privileged information can have an incentive to not trade. Therefore, similarly to the *zeros* measure of Lesmond et al. (1999), idle time can be considered a comprehensive estimate of the degree of liquidity by implicitly including not only the spread, but also commission costs, a portion of the expected price impact costs, and possible opportunity costs of informed traders. This intuition is corroborated by the results of Bandi et al. (2020a) who provide empirical evidence that zero returns (referred to as *stale returns*) are not the mere consequence of market institutional features, such as price rounding. On the contrary, they are related both to the absence of transaction volume and to the magnitude of the bid-ask spread.

We introduce two statistical tests. The first, named *staleness level* test, is designed to validate whether the idle time of a given asset is different from a predefined level, established *a priori* by the researcher. The second, named *staleness equivalence* test, compares the idle times of two assets and assesses whether or not they are statistically indistinguishable.

We study the limiting properties of our test statistics in an asymptotic environment in which the number of observations increases to infinity over a fixed time span. We assume that the underlying price processes evolve as a vector of semimartingales with positive probability of repeated prices. In the staleness level test, we fix the probability of a stale price under the null hypothesis. In the staleness equivalence test, the null is defined as the scenario in which the two assets involved have the same probability of a stale price. For these two tests we show that, under the corresponding null hypotheses, the test statistics converge to asymptotically normal random variables. Moreover, we provide weak convergence results for the estimators of the limiting variances. The alternatives are naturally defined as the scenarios in which the null hypotheses are false. Hence, in the staleness level test, it is possible to statistically assess whether or not an asset has a predefined level of liquidity. In the staleness equivalence test, the null is rejected whenever the two assets involved show two significantly different levels of liquidity.

Considered the relevant role of asset allocation strategies in past and recent studies (see, among others, Golosnoy et al., 2020; Sass & Thös, 2021) we contribute to this literature applying the novel testing procedures to two different diversification problems. In the first application, we show how the staleness equivalence test can be used to obtain groups of equally liquid assets on a daily basis. These groups can be used within a *RnB* estimator of the integrated covariance matrix (Hautsch et al., 2012) to limit the loss of observations due to asynchronous trading and obtain more efficient estimates in large dimension. We evaluate our approach taking the point of view of a mean-variance investor pursuing a volatility timing strategy with a high-dimensional portfolio involving more than one hundred assets. The second application concerns an investor who wants to build a portfolio of assets with a specific liquidity profile. As liquidity is time-varying, the investor needs to control whether the selected assets maintain the desired liquidity level over time. This situation can depict a passive fund willing to replicate the performance of an index, but wants to avoid exposure towards the most illiquid assets of the index. Instead of investing in all the components of the index according to the index weights, the fund can invest only in a subset of highly liquid stocks and then optimize the weights to minimize the tracking error (Jansen & Van Dijk, 2002). Analogously, this situation also suits a hedge fund who wants to gain from an exposure to the most risky assets of an index. Ibbotson et al. (2013) propose to consider illiquidity as an investment style alternative to size, value and momentum. As illiquid assets carry a premium to compensate investors for the liquidation risk, these assets typically provide larger yields. We consider an investor picking stocks with staleness within two pre-specified bounds from a given pool of assets, and allocating funds using an equally-weighted portfolio strategy. The portfolio composition is revisited on a daily basis following the variations in the assets staleness. We show that using a stock selection procedure based on the staleness level test limits the turnover rate and drastically reduces transaction costs compared to a benchmark procedure that simply considers the ranking of idle times.

The reminder of the paper is organized as follows: Section 2 introduces the theoretical framework and presents the main results; Section 3 outlines the Monte Carlo simulations studying the finite sample properties of our tests; Section 4 describes the results of the empirical applications; Section 5 concludes the paper. Mathematical proofs are relegated to the Appendix.

2. Framework

2.1. Setup and Assumptions

Let us consider N assets, and assume that the logarithmic efficient price process of the k-th asset, $Y_t^{(k)}$, follows a Brownian semimartingale

$$Y_t^{(k)} = \int_0^t \mu_s^{(k)} ds + \int_0^t \sigma_s^{(k)} dW_s^{(k)}, \tag{1}$$

where $\mu_s^{(k)}$ denotes an integrable predictable process, $\sigma_s^{(k)}$ is a càdlàg process and $W_s^{(k)}$ is a standard Brownian motion. Further for q, k with $q \neq k$,

$\mathbb{E}[dW_s^{(q)} dW_s^{(k)}] = \rho_{q,k} ds$.

Let $s \in [t, t+1]$ and $t = s_{n,0} < s_{n,1} < \cdots < s_{n,n} = t+1$ be an evenly-spaced partition of $[t, t+1]$, with $\Delta = s_{n,j} - s_{n,j-1}$. Following Bandi et al. (2017, 2020b), we model observed price processes including both an idiosyncratic and a systematic component of staleness. This choice leads to the following reduced-form model for the observed log-price $X^{(k)}_{s_{n,j}}$ on the discrete time grid $\{s_{n,j}\}$

$$\begin{cases} X^{(k)}_{0,n} = Y^{(k)}_{0,n}, \\ X^{(k)}_{j\Delta,n} = (1 - \mathsf{S}_{j,n}) \left(\left(1 - \mathsf{I}^{(k)}_{j,n}\right) Y^{(k)}_{j,n} + \mathsf{I}^{(k)}_{j,n} X^{(k)}_{(j-1)\Delta,n} \right) + \mathsf{S}_{j,n} X^{(k)}_{(j-1)\Delta,n}, \end{cases} \tag{2}$$

where $\mathsf{I}^{(k)}_{j,n}$ and $\mathsf{S}_{j,n}$ are triangular arrays of $\mathcal{F}_j-$measurable Bernoulli variates representing, respectively, the *idiosyncratic* and *systematic* staleness. For more details about the interpretation of the two triangular arrays, $\mathsf{I}^{(k)}_{j,n}$ and $\mathsf{S}_{j,n}$, we refer to Bandi et al. (2020b). Here we limit to say that, should $\mathsf{I}^{(k)}_{j,n} = 1$ then the j-th log-return of the k-th log-price process is zero. If $\mathsf{S}_{j,n} = 1$ the j-th log-asset returns of all assets are set equal to zero. In what follows, we will assume that

$$p^{(k)}_{\mathsf{I},n} \doteq \mathbb{E}\left[\mathsf{I}^{(k)}_{j,n}\right] \xrightarrow{n\to\infty} p^{(k)}_{\mathsf{I}} \in (0,1), \tag{3}$$

$$p_{\mathsf{S},n} \doteq \mathbb{E}\left[\mathsf{S}_{j,n}\right] \xrightarrow{n\to\infty} p_{\mathsf{S}} \in (0,1), \tag{4}$$

where $p^{(k)}_{\mathsf{I},n}$ indicates the idiosyncratic probability of stale prices for asset k at the frequency n, $p_{\mathsf{S},n}$ indicates the systematic probability of stale prices at the frequency n and where $p^{(k)}_{\mathsf{I}}$ and p_{S} denote, respectively, their asymptotic values.

2.2. Testing staleness

To formally define the statistical hypotheses underlying the two tests, we need to rewrite the observed price process in (2) in the following equivalent form

$$X^{(k)}_{j\Delta,n} = (1 - \mathsf{T}^{(k)}_{j\Delta,n}) Y^{(k)}_{j\Delta,n} + \mathsf{T}^{(k)}_{j\Delta,n} X^{(k)}_{(j-1)\Delta,n}, \tag{5}$$

where $\mathsf{T}^{(k)}_{j\Delta,n} \doteq \mathsf{S}_{j\Delta,n} + \mathsf{I}^{(k)}_{j\Delta,n} - \mathsf{S}_{j\Delta,n}\mathsf{I}^{(k)}_{j\Delta,n}$ is a triangular array of $\mathcal{F}_j-$measurable Bernoulli variates representing the *total* staleness for asset k. Accordingly, we define the total staleness probability for asset k at the frequency n as

$$p_{\mathsf{T},n}^{(k)} \doteq \mathbb{E}(\mathsf{T}_{j,n}^{(k)}) \xrightarrow{n\to\infty} p_{\mathsf{T}}^{(k)} \doteq p_{\mathsf{S}} + (1-p_{\mathsf{S}})\, p_{\mathsf{I}}^{(k)}, \tag{6}$$

where $p_{\mathsf{T}}^{(k)}$ indicates the asymptotic total staleness probability for asset k. From (6) one can note that both the systematic and idiosyncratic staleness contribute to the total probability of a stale price for the kth asset. Moreover, given that each asset is equally affected by the systematic component, any difference in total staleness of two given assets can be inputed solely to heterogeneity in the idiosyncratic components. These arguments lead to define the null and the alternative hypotheses of the two tests as follows.

Testing hypothesis 2.1. *Let $\pi \in (0,1)$ be a given real number and let k be an integer $k = 1,...,N$. Let $p_{\mathsf{T}}^{(k)}$ be the total staleness probability of asset k. The null $\mathcal{H}_0^L$ and the alternative $\mathcal{H}_A^L$ hypotheses for the staleness level test are defined, respectively, as*

$$\mathcal{H}_0^L : p_{\mathsf{T}}^{(k)} = \pi, \quad \mathcal{H}_A^L : p_{\mathsf{T}}^{(k)} \neq \pi. \tag{7}$$

Similarly, let $q \neq k$ be an integer, $q = 1,\ldots,N$. The null $\mathcal{H}_0^E$ and the alternative $\mathcal{H}_A^E$ hypotheses for the staleness equivalence test are defined, respectively, as

$$\mathcal{H}_0^E : p_{\mathsf{T}}^{(q)} = p_{\mathsf{T}}^{(k)}, \quad \mathcal{H}_A^E : p_{\mathsf{T}}^{(q)} \neq p_{\mathsf{T}}^{(k)}. \tag{8}$$

Some clarifying observations are in order. A statistical test capable to distinguish between the null and the alternative hypotheses in (7) is required to asses whether a specific asset has a given level of total staleness, no matter its composition in terms of the systematic or the idiosyncratic component. On the other hand, the introduction of a second test capable to discriminate between the null and the alternative hypotheses in (8) is warranted in any situation in which the researcher/investor wants to establish whether or not two assets show statistically comparable levels of total staleness.

Before proceeding to the formal definition of the two statistical tests we need to introduce the following two kernel sample averages:

$$\mathbb{U}_n^{(k)} \doteq \frac{1}{n}\sum_{j=1}^{n} \mathcal{S}\left(\frac{|X_{j\Delta,n}^{(k)} - X_{(j-1)\Delta,n}^{(k)}|}{H_{j,n}^{(k)}}\right), \tag{9}$$

$$\mathbb{M}_n^{(q,k)} \doteq \frac{1}{n}\sum_{j=1}^{n} \mathcal{S}\left(\frac{|X_{j\Delta,n}^{(q)} - X_{(j-1)\Delta,n}^{(q)}|}{H_{j,n}^{(q)}}\right) \mathcal{S}\left(\frac{|X_{j\Delta,n}^{(k)} - X_{(j-1)\Delta,n}^{(k)}|}{H_{j,n}^{(k)}}\right), \tag{10}$$

where $\mathcal{S}(\cdot) : \mathbb{R} \to [0, 1]$ is an integrable function with bounded first derivative in $\mathbb{R}$ and is such that $\mathcal{S}(0) = 1$ ($\mathcal{S}$ is typically called a kernel smoother), and $H_{j,n}^{(k)} = h_n \xi_{j,n}^{(k)}$ is a threshold observed on the evenly-spaced grid, with $\xi_{j,n}^{(k)}$ a bounded positive adapted stochastic process and $h_n \to 0$ as $n \to \infty$. The logic of the estimators $\mathbb{U}_n^{(k)}$ and $\mathbb{M}_n^{(q,k)}$ is clarified in Bandi et al. (2020b). Here we limit to say that, under mild assumptions (listed for convenience in Appendix A.1), Bandi et al. (2020b) show that

$$\mathbb{U}_n^{(k)} \xrightarrow[n\to\infty]{P} p_{\mathsf{S}} + (1 - p_{\mathsf{S}})\, p_{\mathsf{I}}^{(k)}, \tag{11}$$

$$\mathbb{M}_n^{(q,k)} \xrightarrow[n\to\infty]{P} p_{\mathsf{S}} + (1 - p_{\mathsf{S}})\, p_{I}^{(q)}\, p_{\mathsf{I}}^{(k)}. \tag{12}$$

The first convergence result shows that $\mathbb{U}_n^{(k)}$ is a consistent estimator of the total staleness $p_{\mathsf{T}}^{(k)}$, and we use it to define our test statistics and illustrate their asymptotic properties. The second convergence result is needed to estimate the asymptotic distribution of the tests.

Theorem 2.1. *Let $q, k = 1, \ldots, N$ with $q \neq k$ and let $\ell_n^{(k)}$ and $t_n^{(q,k)}$ be the test statistics for the staleness level and staleness equivalence tests defined, respectively, as*

$$\ell_n^{(k)} = \left(\mathbb{U}_n^{(k)} - \pi\right), \tag{13}$$

$$t_n^{(q,k)} = \left(\mathbb{U}_n^{(q)} - \mathbb{U}_n^{(k)}\right). \tag{14}$$

Assume that

$$\lim_{n\to+\infty}\sqrt{n}\left(p_{\mathsf{T},n}^{(k)}-p_{\mathsf{T}}^{(k)}\right)\to 0,\quad k=1,...,N,$$

under the assumptions of Section 2.1, we have that

$$\left(\frac{n}{p_{\mathsf{T},n}^{(k)}\left(1-p_{\mathsf{T},n}^{(k)}\right)}\right)^{\frac{1}{2}}\ell_n^{(k)}\xrightarrow{d}\mathrm{N}(0,1)\qquad(\textit{under }\mathcal{H}_0^L),$$

$$\left(\frac{n}{(1-p_{\mathsf{S},n})\left(p_{\mathsf{I},n}^{(q)}+p_{\mathsf{I},n}^{(k)}-2p_{\mathsf{I},n}^{(q)}p_{\mathsf{I},n}^{(k)}\right)}\right)^{\frac{1}{2}}t_n^{(q,k)}\xrightarrow{d}\mathrm{N}(0,1)\qquad(\textit{under }\mathcal{H}_0^E).$$

Finally, we also have that

$$\left(\frac{n}{p_{\mathsf{T},n}^{(k)}\left(1-p_{\mathsf{T},n}^{(k)}\right)}\right)^{\frac{1}{2}}\ell_n^{(k)}\xrightarrow{p}+\infty\qquad(\textit{under }\mathcal{H}_A^L),$$

$$\left(\frac{n}{(1-p_{\mathsf{S},n})\left(p_{\mathsf{I},n}^{(q)}+p_{\mathsf{I},n}^{(k)}-2p_{\mathsf{I},n}^{(q)}p_{\mathsf{I},n}^{(k)}\right)}\right)^{\frac{1}{2}}t_n^{(q,k)}\xrightarrow{p}+\infty\qquad(\textit{under }\mathcal{H}_A^E).$$

Proof. See Appendix A. □

From Theorem 2.1, we have that the random variables $\ell_n^{(k)}$ and $t_n^{(q,k)}$ are, under their corresponding null hypotheses, asymptotically Gaussian with variances given by, respectively, $\frac{p_{\mathsf{T}}^{(k)}\left(1-p_{\mathsf{T}}^{(k)}\right)}{n}$ and $\frac{(1-p_{\mathsf{S}})\left(p_{\mathsf{I}}^{(q)}+p_{\mathsf{I}}^{(k)}-2p_{\mathsf{I}}^{(q)}p_{\mathsf{I}}^{(k)}\right)}{n}$. On the other side, under the alternatives, they diverge in probability to $+\infty$, delivering an asymptotic unit power. Since, however, the idiosyncratic probabilities $p_{\mathsf{I},n}^{(q)}$ and the systematic probability $p_{\mathsf{S},n}$ are unknown, the variances of the test statistics must be replaced by consistent estimators.

The following corollary illustrates the feasible version of Theorem 2.1 in which suitable estimators of the asymptotic variances are used.

Corollary 2.2. ***Feasible version of the Central Limit Theorem 2.1.***

Let,

$$\widehat{\mathbb{V}}_{\ell_n^{(k)}} \doteq \frac{\mathbb{U}_n^{(k)}\left(1-\mathbb{U}_n^{(k)}\right)}{n},$$
$$\widehat{\mathbb{V}}_{t_n^{(q,k)}} \doteq \frac{\mathbb{U}_n^{(q)}+\mathbb{U}_n^{(k)}-2\mathbb{M}_n^{(q,k)}}{n}.$$

Then, we have that

$$\widehat{\mathbb{V}}_{\ell_n^{(k)}}^{-1/2}\,\ell_n^{(k)} \xrightarrow{d} \mathrm{N}(0,1) \qquad (\textit{under } \mathcal{H}_0^L),$$
$$\widehat{\mathbb{V}}_{t_n^{(q,k)}}^{-1/2}\,t_n^{(q,k)} \xrightarrow{d} \mathrm{N}(0,1) \qquad (\textit{under } \mathcal{H}_0^E).$$

and also that

$$\widehat{\mathbb{V}}_{\ell_n^{(k)}}^{-1/2}\,\ell_n^{(k)} \xrightarrow{p} +\infty \qquad (\textit{under } \mathcal{H}_A^L),$$
$$\widehat{\mathbb{V}}_{t_n^{(q,k)}}^{-1/2}\,t_n^{(q,k)} \xrightarrow{p} +\infty \qquad (\textit{under } \mathcal{H}_A^E).$$

Proof. See Appendix A. □

2.3. Multiple testing problem

Empirical applications might require computing the staleness level and the staleness equivalence tests for a portfolio of assets. Consider the set of N assets, $\left(X^{(1)},\ldots,X^{(N)}\right)$, and let $\mathcal{X}=\{1,...,N\}$ be the collection of the first N integers. Suppose we want to assess whether the assets have the same pre-specified level of staleness, or verify whether they all have the same degree of staleness. To formally define the problem, we specify the hypotheses that need to be tested in a multiple test on the staleness level and staleness equivalence, respectively.

Testing hypothesis 2.2. *Let π be a pre-specified level of total staleness, then the null $\mathcal{H}_0^{ML}$ and the alternative $\mathcal{H}_A^{ML}$ hypothesis for the multiple test on staleness level are*

$$\mathcal{H}_0^{ML}: p_{\mathsf{T}}^{(1)}=\cdots=p_{\mathsf{T}}^{(N)}=\pi \qquad vs \qquad \mathcal{H}_A^{ML}: \exists k\in\mathcal{X}: p_{\mathsf{T}}^{(k)}\neq\pi. \tag{15}$$

Similarly, the null $\mathcal{H}_0^{ME}$ and the alternative $\mathcal{H}_A^{ME}$ hypothesis for the multiple test on staleness equivalence are

$$\mathcal{H}_0^{ME} : p_{\mathsf{T}}^{(1)} = \cdots = p_{\mathsf{T}}^{(N)} \qquad vs \qquad \mathcal{H}_A^{ME} : \exists q, k \in \mathcal{X},\ q \neq k : \ p_{\mathsf{T}}^{(k)} \neq p_{\mathsf{T}}^{(q)}. \tag{16}$$

In the first case, we could simply run the staleness level test for all the assets, and reject $\mathcal{H}_0^{ML}$ if at least one asset rejects the single test. In the second case, we could run the staleness equivalence test across all the pairs, and reject $\mathcal{H}_0^{ME}$ if at least one pair appears statistically different. Since these are multiple testing problems, the size of the tests must be corrected accordingly to avoid spurious over-rejections (Romano et al., 2008). We outline a procedure to do so controlling the family-wise error (FWE) rate.

To test for these hypotheses controlling the FWE rate, we propose the following test statistics,

$$\mathfrak{T}_{\mathcal{X}}^{\mathsf{ML}} = \max_{k \in \mathcal{X}} \left| \frac{\ell_n^{(k)}}{\widehat{\mathbb{V}}^{\frac{1}{2}}_{\ell_n^{(k)}}} \right|, \tag{17}$$

$$\mathfrak{T}_{\mathcal{X}}^{\mathsf{ME}} = \max_{\substack{q,k \in \mathcal{X} \\ q \neq k}} \left| \frac{t_n^{(q,k)}}{\widehat{\mathbb{V}}^{\frac{1}{2}}_{t_n^{(q,k)}}} \right|. \tag{18}$$

Standard results from extreme value theory (Embrechts et al., 1997) show that the normalized maxima of independent Gaussian random variables has a limiting Gumbel distribution, as the number of observations grows to infinity. However, it is clear from Lemma 1 that the test statistics in Equations (17)-(18) are based on dependent random variables.

Lemma 1. *Let $\mathcal{L}_n = \left(\ell_n^{(1)}, \ldots, \ell_n^{(N)}\right)' \in \mathbb{R}^N$ and $\mathcal{T}_n = \left(t_n^{(1,2)}, \ldots, t_n^{(N-1,N)}\right)' \in \mathbb{R}^{\frac{N(N-1)}{2}}$ be the vectors obtained by collecting, respectively, all the test statistics on the staleness level and staleness equivalence. Assume that*

$$\lim_{n \to +\infty} \sqrt{n} \left(p_{\mathsf{T},n}^{(k)} - p_{\mathsf{T}}^{(k)} \right) \to 0, \quad k = 1, ..., N.$$

Then, as $n \to \infty$, it holds that

$$\begin{aligned} \sqrt{n}\mathcal{L}_n &\xrightarrow{d} \mathbf{N}_N\left(0, \Sigma\right) \qquad (\textit{under } \mathcal{H}_0^{ML}), \\ \sqrt{n}\mathcal{T}_n &\xrightarrow{d} \mathbf{N}_{\frac{N(N-1)}{2}}\left(0, \Xi\right) \qquad (\textit{under } \mathcal{H}_0^{ME}), \end{aligned}$$

where $\mathbf{N}_N(\cdot)$ denotes the multivariate N-dimensional Normal distribution, and where the two variance-covariance matrices Σ and Ξ are given, respectively, by

$$\Sigma_{q,k} = \begin{cases} \pi(1-\pi) & q = k \\ p_{\mathsf{S}} + (1-p_{\mathsf{S}})p_{\mathsf{I}}^{(q)}p_{\mathsf{I}}^{(k)} - \pi^2 & q \neq k \end{cases} \tag{19}$$

$$\Xi_{(h,k),(r,q)} = \begin{cases} (1-p_{\mathsf{S}})\left(p_{\mathsf{I}}^{(h)} - 2p_{\mathsf{I}}^{(h)}p_{\mathsf{I}}^{(k)} + p_{\mathsf{I}}^{(k)}\right) & h = r,\ k = q \\ (1-p_{\mathsf{S}})\left(p_{\mathsf{I}}^{(h)} - p_{\mathsf{I}}^{(h)}p_{\mathsf{I}}^{(q)} - p_{\mathsf{I}}^{(k)}p_{\mathsf{I}}^{(h)} + p_{\mathsf{I}}^{(k)}p_{\mathsf{I}}^{(q)}\right) & h = r,\ k \neq q \\ 0 & h \neq r,\ k \neq q \end{cases}$$

Proof. See Appendix A. □

To obtain the distribution of $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ML}}$ and $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ME}}$, we rely on the results of Arellano-Valle & Genton (2008), who derive the exact distribution for the maximum of dependent Gaussian random variables.

Lemma 2. *Consider, for a fixed i, the following partition of the asymptotic covariance matrix of $\mathcal{L}_n$ in Lemma 1,*

$$\Sigma = \begin{pmatrix} \Sigma_{-i-i} & \Sigma_{-ii} \\ \Sigma_{i-i} & \Sigma_{ii} \end{pmatrix},$$

where the subscript $-i$ indicates the removal of the ith row or column of Σ and Σ_{ii} denotes the element of Σ on the ith row of the ith column. The p.d.f. of $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ML}}$ can be written as

$$f_{\mathfrak{T}_{\mathcal{X}}^{\mathsf{ML}}}(x) = \sum_{i=1}^{N} \phi\left(0, \Sigma_{ii}\right) \Phi_{N-1}\left(x\frac{\Sigma_{-ii}}{\Sigma_{ii}}, \Sigma_{-i-i} - \frac{\Sigma_{-ii}\Sigma_{-ii}'}{\Sigma_{ii}}\right), \tag{20}$$

with $\phi(\cdot)$ *and* $\Phi_N(\cdot)$ *denoting the marginal Gaussian density and the* N*-variate Gaussian cumulative distribution, respectively. Analogous result can be obtained for* $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ME}}$*.*

Proof. The result follows from Lemma 1 and Corollary 4 of Arellano-Valle & Genton (2008) □

The computational cost to obtain the exact limiting density in Equation (20) is quite high when the number of assets is large. As our empirical application involves more than 150 assets, we compute the critical values of the test statistics using Monte Carlo simulations instead of the theoretical result in Equation (20). We compare exact and simulated critical values in a small setting and find that differences are negligible.

The two test statistics $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ML}}$ and $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ME}}$, while controlling for the FWE rate, do not identify the source of rejection. Consider, for example, the case of $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ML}}$: the rejection of the null does not say which assets, among those included in $\mathcal{X}$, have an estimated level of total staleness statistically different from the pre-defined level π. A similar argument applies for $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ML}}$. To solve this issue, we propose the following algorithm: first we run the multiple test (either $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ML}}$ or $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ME}}$) using the whole set of indexes $\mathcal{X}$. Then, if the multiple null hypothesis is rejected, we remove from $\mathcal{X}$ the asset with the highest value of the test statistic, obtaining a new set $\mathcal{X}' \subset \mathcal{X}$. We repeat the first step on the new set $\mathcal{X}'$ and we iterate the procedure until the null is not rejected anymore. At the end of the procedure the original set $\mathcal{X}$ is split into two non-overlapping sets $\mathcal{X} = \mathcal{X}_0 \bigcup \mathcal{X}_1$. In the case of $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ML}}$ (resp. $\mathfrak{T}_{\mathcal{X}}^{\mathsf{ME}}$) the set $\mathcal{X}_1$ contains all the assets (resp. all the pairs of assets) for which the hypothesis $\mathcal{H}_0^L$ (resp. $\mathcal{H}_0^E$) is rejected.

3. Simulations

This section studies the finite sample sizes and powers of the tests outlined in Sections 2.2 and 2.3. We consider a set of N assets and assume that the efficient log-price $Y^{(k)}$ of the kth asset obeys the following factor structure,

$$\begin{aligned} dF_t &= \sigma_F \, dW_t^F, \\ dY_t^{(k)} &= \beta^{(k)} \, dF_t + \sigma_{(k)} \, dW_t^{(k)}, \end{aligned} \tag{21}$$

where F_t is a factor common to all the N assets, $\left(dW_t^F, dW_t^{(1)}, \ldots, dW_t^{(N)}\right)$ are independent standard Brownian motions, and $\beta^{(k)}$ is a factor loading drawn from a normal distribution, $N\left(\overline{\beta}, \sigma_\beta\right)$, to consider an heterogeneous impact of the factor on the assets. We generate a trading day of six hours setting $\sigma_F = 0.01$, $\sigma_{(k)} = 0.01$, $\overline{\beta} = 0.5$ and $\sigma_\beta = 0.45$. Then, we consider sampling schemes at the frequencies $\Delta = (1, 10, 30, 60, 300)$ seconds and obtain the observed price process on these grids as

$$\begin{cases} X_{0,n}^{(k)} = Y_{0,n}^{(k)}, \\ X_{j,n}^{(k)} = (1 - \mathsf{S}_{j,n}) \left(\left(1 - \mathsf{I}_{j,n}^{(k)}\right) Y_{j,n}^{(k)} + \mathsf{I}_{j,n}^{(k)} X_{(j-1),n}^{(k)} \right) + \mathsf{S}_{j,n} X_{(j-1),n}^{(k)}, \end{cases} \tag{22}$$

where $\left(\mathsf{S}_{j,n}, B_{j,n}^{(1)}, \ldots, B_{j,n}^{(N)}\right)$ are independent Bernoulli random variables with $\mathbb{E}\left[\mathsf{S}_{j,n}\right] = p_{\mathsf{S},n}$ and $\mathbb{E}\left[B_{j,n}^{(k)}\right] = p_{\mathsf{I},n}^{(k)}$, $k = \{1, \ldots, N\}$. To obtain the finite sample staleness probabilities, we consider the following scaling laws

$$\begin{aligned} p_{\mathsf{S},n} &= p_{\mathsf{S}} \left(1 - \exp\left(-0.001 \, n\right)\right), \\ p_{\mathsf{I},n}^{(k)} &= p_{\mathsf{I}}^{(k)} \left(1 - \exp\left(-0.001 \, n\right)\right) \quad \forall k \in \{1, \ldots, N\}, \end{aligned}$$

where $n = \frac{1}{\Delta}$ and $\left(p_{\mathsf{S}}, p_I^{(1)}, \ldots, p_I^{(N)}\right)$ are the limiting staleness probabilities. This scaling law simply reflects the empirical evidence that the higher the sampling frequency the higher the likelihood of a stale price (Bandi et al., 2020b).

For each frequency, the kernel averages $\mathbb{U}_n^{(k)}$ and $\mathbb{M}_n^{(q,k)}$ in Equations (9)-(10) are computed replacing the kernel $\mathcal{S}(\cdot)$ with the indicator function for zero returns, in formula

$$\mathbb{U}_n^{(k)} = \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{\left\{X^{(k)}_{j\Delta,n} - X^{(k)}_{(j-1)\Delta,n} = 0\right\}}, \tag{23}$$

$$\mathbb{M}_n^{(q,k)} = \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{\left\{X^{(q)}_{j\Delta,n} - X^{(q)}_{(j-1)\Delta,n} = 0\right\}} \mathbb{1}_{\left\{X^{(k)}_{j\Delta,n} - X^{(k)}_{(j-1)\Delta,n} = 0\right\}}, \tag{24}$$

whence $\mathbb{U}_n^{(k)}$ and $\mathbb{M}_n^{(q,k)}$ represent, respectively, the percentage of zero returns for the asset k and the percentage of simultaneous zero returns for asset q and k.

The adoption, in the definitions (23) and (24), of an indicator function in place of a kernel smoother $\mathcal{S}(\cdot)$, brings the double advantage of I) a less nuanced interpretation of the moment estimators (percentages of zero returns) and II) of a smaller estimator bias in finite sample. Kernel-based estimators are key only when a large number of assets are simultaneously involved. In this context the bivariate moment estimator $\mathbb{M}_n^{(q,k)}$ must be replaced with its N-variate (for a generic number of assets N) counterpart, which is naturally defined as

$$\mathbb{M}_n^{(N)} = \frac{1}{n}\sum_{j=1}^{n}\prod_{k=1}^{N} \mathbb{1}_{\left\{X^{(k)}_{j\Delta,n} - X^{(k)}_{(j-1)\Delta,n} = 0\right\}}. \tag{25}$$

As a matter of fact, as the number of assets involved increases, the across-assets product of indicator functions that appear in Equation (25) becomes, in finite sample, prone to produce a zero-valued estimator even in presence of high level of systematic staleness (it is enough that one out of the N assets involved has a non-zero return to have a product identically equal to zero, no matter how large, but finite, N is). As a consequence, for finite but large N, the choice of a finite (small) threshold (as in Bandi et al., 2020b) is unavoidable. This choice is paid for in terms of a positive bias in the estimation of price staleness: a non-null threshold is deemed to include, in the estimation, small fluctuations due to the efficient price volatility, inflating the estimation. The case analyzed here is different. Since no more than two assets are involved at a time, the choice of a kernel smoother would be, in finite sample, suboptimal. This said, both definitions (either with indicator functions or kernels) give the same results asymptotically (i.e., when the number of observations n and the number of assets N diverge simultaneously), see Bandi et al. (2020b) for more details. We decided to frame the theory in terms of kernel functions (for sake of generality) and to adopt indicator functions in the empirical application (for the finite sample advantages just discussed).

3.1. Testing the staleness level

We assess the finite sample properties of the test on the staleness level, assuming a single stock, i.e. $N = 1$. Given that we simulate only one process, it is not necessary to distinguish between idiosyncratic and systematic staleness, hence we consider three scenarios characterized by three different levels of (asymptotic) total staleness probability, more specifically $p_{\mathsf{T}} \in \{0.1, 0.3, 0.5\}$.

To verify that the staleness level test, associated to the statistic defined in Equation (13), is properly sized we implement the test with $\pi = p_{\mathsf{T}}$. Similarly, to check the power, the test is implemented inputing a $\pi \neq p_{\mathsf{T}}$, in particular we choose $\pi = p_{\mathsf{T}} + \pi_{\Delta}$, where $\pi_{\Delta} \in \{0.01, 0.025, 0.05\}$.

Table 1 displays the power and size of the test statistic $\ell_n^{(k)}$ for different significance levels α, and several sampling frequencies, $\Delta = \{1, 10, 30, 60, 300\}$ seconds, over 1000 replications. The test is properly sized at high frequency, but one can note that distortions appear at moderate frequencies. At the 5-minute sampling frequency, the test exhibits a high number of over-rejections, and this evidence is stronger the greater p_{T}. This finite sample distortion emerges because of the discrepancy between $p_{\mathsf{T},n}$ and p_{T}, and clearly vanishes as n increases because $p_{\mathsf{T},n}$ converges to p_{T}. At the same time, the test is very powerful across all the sampling frequencies and for all the values of π_{Δ}.

To study the properties of the multiple test statistic on the staleness level in Equation (17), we consider a realistic setting where the N stocks are divided in G groups with different degrees of staleness: this means that each group contains stocks exhibiting the same probability of a stale price. We consider three dimensions, $N = \{80, 120, 160\}$, four numbers of groups, $G = \{1, 2, 4, 8\}$, and three asymptotic systematic staleness probabilities $p_{\mathsf{S}} = \{0.005, 0.01, 0.05\}$. Recall that $\mathcal{X}$ is composed of two subsets, i.e. $\mathcal{X} = \{\mathcal{X}_0, \mathcal{X}_A\}$, where $\mathcal{X}_0$ and $\mathcal{X}_A$ contain the indices k of the assets which do and do not satisfy the null hypothesis, respectively. When $G = 1$, we have that $\mathcal{X} \equiv \mathcal{X}_0$, and we can assess the ability to control the FWE rate. The asymptotic idiosyncratic staleness probability is set to $p_{\mathsf{I}}^{(k)} = 0.1$, and the staleness probability under $\mathcal{H}_0^{ML}$ is $\pi = p_{\mathsf{S}} + (1 - p_{\mathsf{S}})p_{\mathsf{I}}^{(k)}$. The upper panel of Table 2 shows the FWE rate at level $\alpha = 0.05$ over 100 replications. These results are similar to those observed for the size of staleness level test (Table 1), suggesting that small sample effects carry over to the multiple testing problem. To assess the power of the test, we consider $G = \{2, 4, 8\}$. We spread the idiosyncratic staleness probabilities of the assets over an equally-spaced grid within the interval $\overline{p}^L = 0.1$ and $\overline{p}^H = 0.5$. Assets in the first group have idiosyncratic staleness probability equal to $\overline{p}^L$, while the idiosyncratic staleness probability of the assets in the g-th group is equal to $\overline{p}^L + g\frac{\overline{p}^H - \overline{p}^L}{G}$, with $g \in \{2, \ldots, G\}$. We assume that the assets in the first group belong to $\mathcal{X}_0$, i.e. $\pi = p_{\mathsf{S}} + (1 - p_{\mathsf{S}})\overline{p}^L$, while those in the others $G-1$ groups belong to $\mathcal{X}_A$. We assess the ability of rejecting the null hypothesis counting the number of assets correctly allocated to the subset $\mathcal{X}_A$. Results in Table 2 confirm that the test is very powerful up to the 1-minute sampling frequency.

Table 1: Power and Size for the test on the staleness level

π_Δ	Δ (sec.)	$p_{\mathsf{T}} = 0.1$			$p_{\mathsf{T}} = 0.3$			$p_{\mathsf{T}} = 0.5$		
α		0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
0	1	0.115	0.054	0.011	0.097	0.052	0.014	0.083	0.049	0.013
	10	0.120	0.064	0.018	0.087	0.042	0.005	0.094	0.042	0.010
	30	0.118	0.062	0.015	0.106	0.043	0.009	0.104	0.054	0.011
	60	0.149	0.108	0.039	0.203	0.117	0.025	0.228	0.119	0.042
	300	0.682	0.682	0.473	0.943	0.906	0.784	0.999	0.994	0.982
0.01	1	1.000	0.999	0.993	0.969	0.939	0.794	0.941	0.895	0.726
	10	0.563	0.423	0.215	0.418	0.281	0.104	0.371	0.232	0.074
	30	0.340	0.254	0.098	0.267	0.166	0.065	0.256	0.168	0.042
	60	0.306	0.255	0.118	0.325	0.203	0.074	0.342	0.226	0.092
	300	0.812	0.656	0.467	0.978	0.957	0.855	1.000	0.999	0.983
0.025	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	10	0.987	0.969	0.903	0.877	0.804	0.566	0.828	0.720	0.472
	30	0.791	0.701	0.464	0.567	0.423	0.217	0.545	0.404	0.187
	60	0.657	0.516	0.313	0.500	0.373	0.176	0.575	0.410	0.189
	300	0.915	0.813	0.651	0.983	0.969	0.909	1.000	0.998	0.985
0.05	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	10	1.000	1.000	1.000	1.000	0.999	0.992	0.999	0.999	0.992
	30	0.998	0.994	0.975	0.949	0.904	0.741	0.912	0.850	0.630
	60	0.979	0.948	0.813	0.863	0.775	0.519	0.885	0.772	0.535
	300	0.962	0.911	0.820	0.997	0.987	0.953	1.000	1.000	0.997

This table shows the rejection frequencies of the test on the staleness level $\ell_n^{(k)}$ over 1000 replications. We assume three different levels of total staleness probability p_{T}. Δ refers to the sampling frequency and α is the nominal size of the test.

Table 2: FWE rate and Power of the Multiple Test on Staleness Level

	G	Δ (sec.)	$N=80$			$N=120$			$N=160$		
p_{S}			0.001	0.005	0.05	0.001	0.005	0.05	0.001	0.005	0.05
	1	1	0.060	0.040	0.020	0.040	0.060	0.080	0.090	0.060	0.110
		10	0.060	0.070	0.080	0.080	0.090	0.010	0.120	0.070	0.070
		30	0.080	0.130	0.070	0.120	0.080	0.100	0.090	0.100	0.130
		60	0.220	0.250	0.180	0.210	0.250	0.100	0.360	0.330	0.150
		300	1.000	1.000	0.930	1.000	1.000	0.900	1.000	1.000	0.910
	2	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		30	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		60	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		300	0.346	0.304	0.141	0.265	0.260	0.114	0.255	0.231	0.094
	4	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		30	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		60	1.000	1.000	0.998	0.999	1.000	0.996	0.999	0.999	0.995
		300	0.114	0.103	0.054	0.096	0.098	0.036	0.092	0.078	0.037
	8	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		30	0.992	0.990	0.971	0.990	0.987	0.964	0.985	0.984	0.961
		60	0.912	0.910	0.885	0.907	0.902	0.876	0.899	0.897	0.875
		300	0.073	0.073	0.040	0.076	0.061	0.030	0.064	0.055	0.027

This table shows the finite sample properties of the multiple test on the staleness level over 100 replications. The upper panel ($G=1$) shows the FWE rate. The other panels ($G=\{2,4,8\}$) show the power of the test, measured as the fraction of assets correctly assigned to the set under the alternative $\mathcal{X}_A$.

3.2. Testing staleness equivalence

We assess the finite sample properties of the staleness equivalence test in Equation (14). For this purpose, we consider a set of two assets, i.e. $N = 2$. We assume that the first asset has an idiosyncratic staleness probability $p_{\mathsf{I}}^{(1)} = 0.30$, while for the second asset, we consider four different scenarios, i.e. $p_{\mathsf{I}}^{(2)} = \{0.30, 0.35, 0.40, 0.50\}$. The case $p_{\mathsf{I}}^{(1)} = p_{\mathsf{I}}^{(2)}$ is used to study the size of the test. Table 3 displays the power and size of the test statistic at different significance level α, for various level of the systematic staleness probability p_{S}, and for different sampling frequencies, $\Delta = \{1, 10, 30, 60, 300\}$ seconds, over 1000 replications. The results in Table 3 highlight that the test is properly sized and the various scenarios do not lead to significant size distortions. At the same time, the test is already quite powerful for $p_{\mathsf{I}}^{(2)} = 0.35$ at the 1-minute sampling frequency, though less powerful than the staleness level test. The lower power is a consequence of the fact that the staleness equivalence test requires the three estimators $\mathbb{U}_n^{(1)}$, $\mathbb{U}_n^{(2)}$, and $\mathbb{M}_n^{(1,2)}$ instead of one ($\mathbb{U}_n^{(1)}$), slowing convergence to normality in finite samples. Finally, we note that the level of the systematic component has hardly any impact on the power of the test.

To study the properties of the multiple testing procedure of staleness equivalence, we create a realistic setting where the N stocks are divided in G equally-sized groups. Stocks in the same group exhibit the same degree of staleness, but the latter differs across the groups. Similarly to the simulations for the multiple staleness level test, we set the minimum and maximum idiosyncratic staleness probability to $\overline{p}^L = 0.1$ and $\overline{p}^H = 0.9$, and let the intra-group staleness be $\frac{\overline{p}^H - \overline{p}^L}{G}$ distance apart. We consider several scenarios where $N = \{80, 120, 160\}$ and $G = \{1, 2, 4, 8\}$. When $G = 1$, we can assess the ability to control the FWE rate. When $G = \{2, 4, 8\}$, we assume that the assets in the first group belong to $\mathcal{X}_0$ and the others to $\mathcal{X}_A$, and evaluate the power of the test counting the number of assets correctly allocated to the subset $\mathcal{X}_A$. Table 4 reports the size and power of the multiple test on staleness equivalence at the significance level $\alpha = 0.05$ and at the sampling frequency $\Delta = 60$ seconds (these values will be used in the empirical applications), over 100 replications. The results highlight that the multiple test on staleness equivalence is properly sized for all levels of the systematic staleness probability, and for all the number of assets considered. As to the power, we highlight that the test is extremely powerful for all the considered cases.

Table 3: Power and Size of the Staleness Equivalence test

$p_{\mathsf{I}}^{(2)}$	Δ (sec.)	$p_{\mathsf{S}} = 0.005$			$p_{\mathsf{S}} = 0.01$			$p_{\mathsf{S}} = 0.05$		
α		0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
0.30	1	0.115	0.058	0.010	0.086	0.039	0.007	0.105	0.054	0.013
	10	0.093	0.052	0.012	0.099	0.051	0.013	0.110	0.053	0.017
	30	0.097	0.043	0.008	0.085	0.042	0.007	0.093	0.045	0.003
	60	0.099	0.055	0.009	0.104	0.049	0.010	0.091	0.044	0.007
	300	0.095	0.053	0.006	0.108	0.056	0.010	0.113	0.057	0.010
0.35	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	10	0.986	0.965	0.876	0.987	0.967	0.885	0.989	0.962	0.869
	30	0.764	0.629	0.374	0.767	0.640	0.384	0.785	0.654	0.378
	60	0.557	0.409	0.183	0.555	0.403	0.187	0.507	0.374	0.176
	300	0.190	0.107	0.028	0.182	0.090	0.017	0.199	0.101	0.024
0.40	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
	30	0.994	0.990	0.950	0.996	0.988	0.948	0.994	0.982	0.932
	60	0.927	0.867	0.672	0.931	0.881	0.678	0.934	0.850	0.619
	300	0.337	0.232	0.068	0.295	0.184	0.044	0.314	0.201	0.044
0.50	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	30	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	60	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000	0.998
	300	0.564	0.409	0.192	0.576	0.442	0.206	0.588	0.430	0.191

This table shows the rejection frequencies of the test on the staleness equivalence $t_n^{(q,k)}$ over 1000 replications. We assume three different levels of the systematic staleness probability p_{S}, and set $p_{\mathsf{I}}^{(1)} = 0.30$. Δ refers to the sampling frequency and α is the nominal size of the test.

Table 4: FWE rate and Power of Multiple Test on Staleness Equivalence

G	$N = 80$			$N = 120$			$N = 160$		
p_{S}	0.001	0.005	0.05	0.001	0.005	0.05	0.001	0.005	0.05
1	0.050	0.030	0.050	0.060	0.070	0.020	0.080	0.080	0.070
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
8	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999

FWE rate and power of the multiple test on staleness equivalence for three different levels of the systematic staleness probability p_{S}. G is the number of groups considered in the simulation and N is the number of simulated assets. The row with $G = 1$ reports the results for the FWE rate, whereas the other ones refer to the power. The test is applied with a confidence level of $\alpha = 0.05$ and $\Delta = 60$ seconds.

4. Empirical applications

This section presents two applications of the testing procedures outlined above.

4.1. Data

Our empirical analysis is based on high frequency data for $N = 152$ companies from the NYSE, ranging from January 2006 until December 2014 ($T = 2183$ days). We observe prices at the 1-minute frequency. Idle times $\mathbb{U}_n^{(k)}$ and $\mathbb{M}_n^{(q,k)}$ in Equations (9)-(10) are computed at one-minute frequency as well. As in the Monte Carlo simulations, we replace the kernels with the indicator functions, see Equations (23)-(24).

4.2. Volatility timing with optimal staleness groups

We consider an investor pursuing a *volatility timing* asset allocation strategy (Fleming et al., 2001, 2003) with a vast dimensional portfolio, re-balancing on a daily basis. On any day t, the investor allocates funds into N risky assets solving the global minimum variance portfolio problem,

$$\min_{\mathbf{w}_t} \mathbf{w}_t' \mathbf{\Sigma}_t \mathbf{w}_t,$$
$$s.t. \, \mathbf{w}_t' \boldsymbol{\iota} = 1,$$

where $\mathbf{w}_t$ is a $N \times 1$ vector of portfolio weights, $\mathbf{\Sigma}_t$ is the $N \times N$ integrated covariance matrix on day t, $\boldsymbol{\iota}$ is a $N \times 1$ vector of ones. When $\mathbf{\Sigma}_t$ is positive definite, the optimal portfolio allocation has closed-form solution

$$\mathbf{w}_t = \frac{\mathbf{\Sigma}_t^{-1} \boldsymbol{\iota}}{\boldsymbol{\iota}' \mathbf{\Sigma}_t^{-1} \boldsymbol{\iota}}. \tag{26}$$

Our aim is to obtain accurate estimates of the latent daily integrated covariance matrix $\mathbf{\Sigma}_t$ using high-frequency data. This task can be very challenging due to the asyncronicity characterizing stock prices sampled at high frequency, which induces a downward bias in the estimates of the daily covariance. A typical solution is to resort to *refresh time* sampling, i.e. create a discrete time grid where price updates are synchronized, and use the synchronized prices to estimate the integrated covariance. Although this strategy removes the bias, it forces to discard a large amount of observations, particularly if one of the stocks in portfolio is extremely illiquid. To limit this problem, Hautsch et al. (2012) propose a *Regularizing and Blocking* (RnB) estimator of the integrated covariance, working in two steps: first, stocks with homogeneous liquidity are grouped together, and the daily covariance of each group is estimated. They suggest to use an estimator robust to the microstructure noise, such as the *Multivariate Realized Kernel* (MRK) of Barndorff-Nielsen et al. (2011); second, the covariances from the sub-groups are combined to obtain the entire covariance matrix, and the latter is regularized in order to ensure it is positive definite.

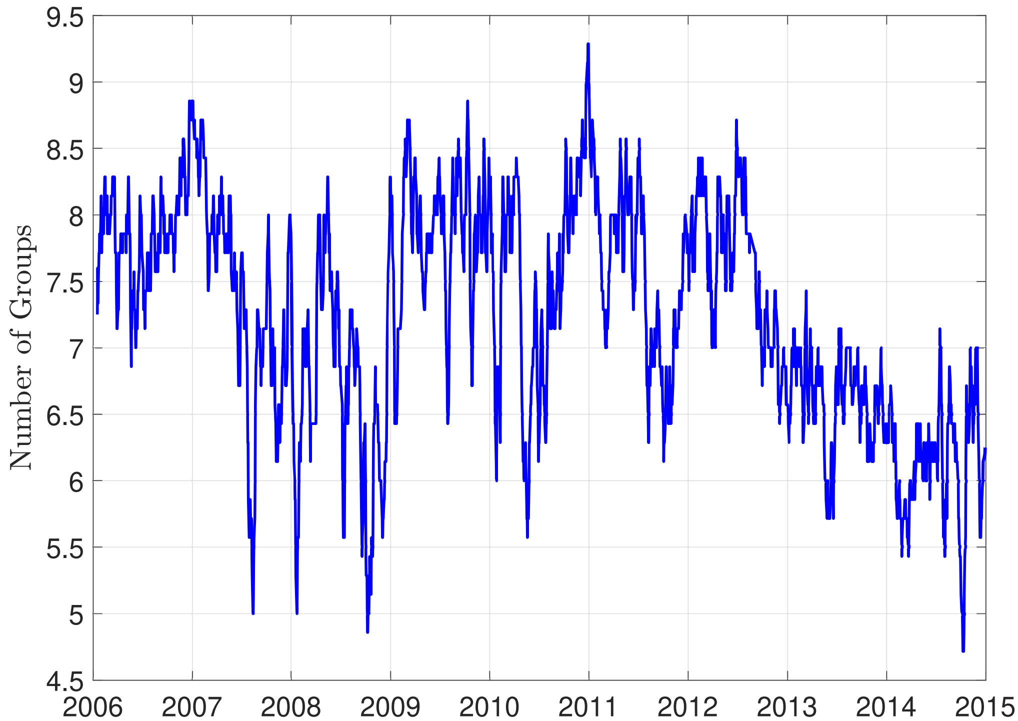
The approach of Hautsch et al. (2012) is based on a k-means algorithm applied to the number of daily transactions. Their procedure forces to keep the number of groups and the groups composition fixed throughout the days of the sample. We propose a flexible procedure to obtain daily groups of equally-liquid assets. We sequentially implement the multiple staleness equivalence testing strategy (outlined in Section 2.3). At each iteration, the stock with the highest staleness is removed from the group (which, initially, contains all the assets) until the null hypothesis of staleness equivalence among the assets is rejected. Once a group is established, a new sequence of iterations starts. The multiple staleness equivalence test is now performed on the assets removed from the previous groups. The procedure is iterated until the null hypothesis cannot be rejected anymore. Thus, when the iteration is ended, all the assets are assigned to a group. This procedure allows to have data-driven group size and structure that change every day. We refer to these groups as *optimal staleness groups*. As in Hautsch et al. (2012), we estimate the daily integrated covariance using the MRK estimator of Barndorff-Nielsen et al. (2011) on the optimal staleness groups, stacking the sub-group covariances and regularizing the entire covariance matrix. We name this the *Staleness Group* (SG) estimator.

The relevance of accounting for a dynamic group structure is depicted in Figure 1, which shows the monthly average number of identified groups using the procedure based on the staleness equivalence test. The plot shows that the number of groups changes over time, reaching the minima during periods of turmoil. This suggests that when the trading activity hastens, the number of classes shrinks due to a stronger market concentration. On the other hand, the number of identified groups is higher over days characterized by a more regular trading activity. Besides the dynamic number of groups, we emphasize the importance of letting also the group structure to vary over time. Figure 2 shows the monthly idle time ranking of the stocks described in Section 4.1. The plot highlights how stocks tend to change their position in the ranking, suggesting that assuming a constant group structure over time is restrictive.

To assess the merit of our SG estimator for portfolio optimization, we rely on the approach of Engle & Colacito (2006). We consider several alternative estimators of the integrated covariance, and use Equation (26) to obtain global minimum variance portfolio solutions with each one of them. The estimator returning the lowest portfolio variance is the most accurate. The analysis is performed in-sample so that we do not need to specify any time series model to obtain forecasts. We consider the 5-minute realized volatility of the portfolio returns as a measure of portfolio volatility $\sigma_{p,t}$. As competitors of our SG estimator, we select the MRK of Barndorff-Nielsen et al. (2011), and the RnB estimator of Hautsch et al. (2012) using 2, 4 and 8 blocks. As these estimators are not positive definite, we need a regularization method to ensure that this property holds. We rely on the eigenvalue cleaning (EC) approach used in Hautsch et al. (2012) and on the factor model regularization using a number of factors equal to one (F1) and three (F3) as in Hautsch et al. (2015).

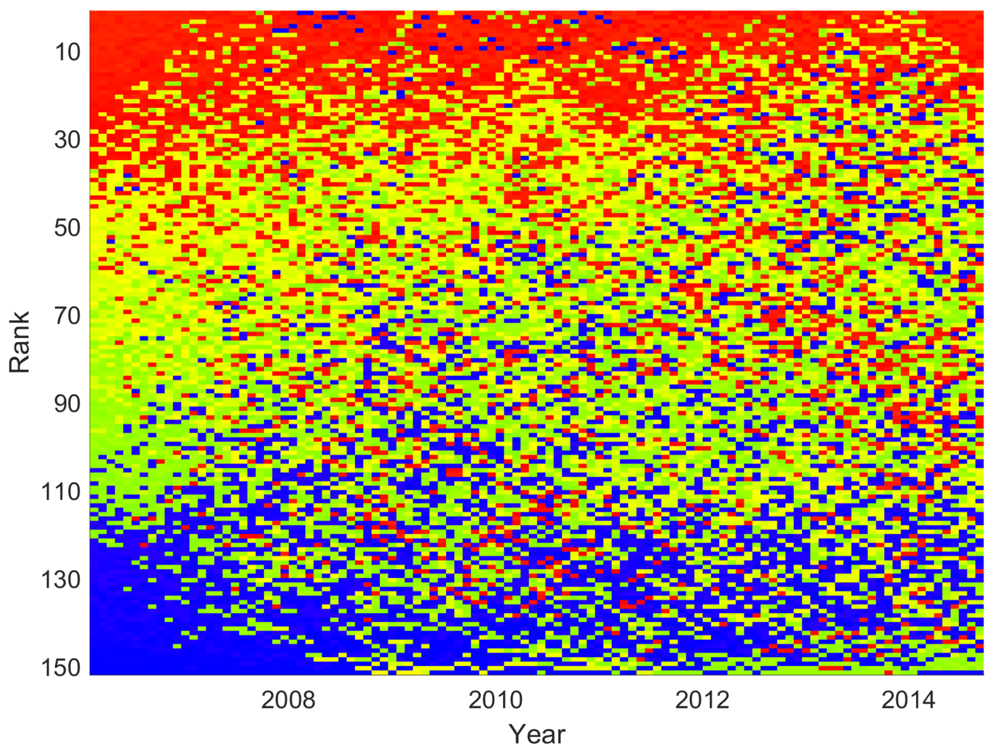
Table 5 displays the average annualized portfolio volatility obtained with the different estimators and regularization methods on the 152 companies described in Section 4.1. For each regularization method, the SG estimator returns the lowest portfolio variance. To formally determine whether the superior performance of our SG estimator is statistically significant, we apply the Model Confidence Set (MCS) of Hansen et al. (2011) to the portfolio volatility estimates obtained from each regularization method. The MCS uses a bootstrap procedure to identify the (sub)set of "best forecasting" models with a given confidence level. In the implementation, we use 10000 bootstrap samples obtained from a block bootstrap with block size of 20 observations, and build the optimal subset at the 90% confidence level using the maximum difference criterion. Results in Table 5 shows that, for each regularization method, the estimate obtained with the SG estimator is the only one included in the Model Confidence Set (MCS) of Hansen et al. (2011). Moreover, when we jointly consider all the estimators and regularization methods, the SG estimator regularized with EC method is the overall winner.

Figure 1: Average number of identified groups



This graph shows the monthly average number of identified groups using the grouping procedure based on the staleness equivalence test.

Figure 2: Idle time heatmap



This plot sorts assets based on their idle time computed on a monthly basis. We assign a specific color to each stock depending on its position in the ranking in the first month. Tracking the firms over time clarifies to what extent stocks that used to be illiquid become more liquid and viceversa.

Table 5: Volatility of different IC estimators

Estimator	Regularization Method		
	EC	F1	F3
MRK	0.0722	0.0711	0.0728
SG	**0.0578**	0.0652	0.0716
RnB2	0.0644	0.0789	0.1428
RnB4	0.0615	0.0716	0.1179
RnB8	0.0596	0.0849	0.1047

This table shows the average annualized portfolio volatility obtained with the different estimators and regularization methods. Underlined numbers highlight which estimator is in the Model Confidence Set for a given regularization method. Bold numbers denote that an estimator is part of the Model Confidence Set considering all regularization methods.

4.3. Staleness-based asset allocation

Liquidity is a relevant risk factor in the asset allocation decisions of investors. Some investors prefer to allocate funds towards highly liquid assets to reduce the impact of transaction costs on the portfolio performance. For instance, large passive funds trying to replicate the performance of an equity index may prefer to hold only a subset of the index components instead of the entire index, in order to minimize the costs associated to the most illiquid assets (Jansen & Van Dijk, 2002). Other investors, such as hedge funds, may prefer to expose themselves to highly illiquid assets, to benefit from the liquidity risk premium (Ibbotson et al., 2013). Both examples suggest that investors can benefit from a stock picking procedure allowing them to select stocks with a specific liquidity profile, taking also into account the dynamic nature of liquidity. We meet the investor's need developing a simple procedure building portfolios with a pre-specified degree of staleness.

We consider an investor who can allocate funds in the 152 companies described in Section 4.1, but is willing to invest only in those stocks with total staleness $p_{\mathsf{T}}^{(k)}$ within a given staleness *fork*, i.e. $p_{\mathsf{T}}^{(k)} \in (\pi_L, \pi^U)$, with $0 < \pi_L < \pi^U < 1$. On any given day t, the investor picks the stocks, and then *naively* distributes its funds over the selected stocks following a simple $\frac{1}{n}$ rule. Nothing prevents the investor to optimize the portfolio weights solving the Markowitz problem but we prefer to adopt a $\frac{1}{n}$ rule to assess the merit of our stock picking procedure. Formally, let $\omega_{k,t}$ be the portfolio weight of asset k on day t, then

$$\omega_{k,t} = \begin{cases} \frac{1}{n_t^{\pi}} & \pi_L \leq p_{\mathsf{T},t}^{(k)} \leq \pi^U \\ 0 & p_{\mathsf{T},t}^{(k)} < \pi_L \text{ or } p_{\mathsf{T},t}^{(k)} > \pi^U \end{cases},$$

where $p_{\mathsf{T},t}^{(k)}$ is the total staleness probability of the asset k on day t and n_t^{π} denotes the number of assets whose total staleness falls in the interval (π_L, π^U) on day t.

This task would be easy if we could measure the probability of a stale price without errors, but we can only obtain a noisy estimate of it using idle time. As the latter is naturally characterized by some degree of uncertainty, we may have assets that belong to the interval (π_L, π^U), but whose idle times lay outside of it. Spurious exclusion of these assets limits portfolio diversification and inevitably increases the portfolio turnover, thus raising the transaction costs. To avoid this issue, we use the test of staleness level to select the assets to include in the investor's portfolio. On any given day, for each asset k, we adopt the following selection rule:

- If $\pi_L < \mathbb{U}^{(k)} < \pi^U$, we include the kth asset in the portfolio.
- If $\mathbb{U}^{(k)} < \pi_L$, then we test the null hypothesis $\mathcal{H}_0 : \mathbb{U}^{(k)} = \pi_L$ against the alternative $\mathcal{H}_A : \mathbb{U}^{(k)} < \pi_L$. Rejection of $\mathcal{H}_0$ favors the exclusion of the kth asset from the portfolio.
- If $\mathbb{U}^{(k)} > \pi^U$, then we test the null hypothesis $\mathcal{H}_0 : \mathbb{U}^{(k)} = \pi^U$ against the alternative $\mathcal{H}_A : \mathbb{U}^{(k)} > \pi^U$. Rejection of $\mathcal{H}_0$ favors the exclusion of the kth asset from the portfolio.

As we need to test all the assets laying outside the staleness interval, we account for the multiple testing problem using the results of Section 2.3.

We use this procedure to build a portfolio Π_P with staleness between $\pi_L = 0.5$ and $\pi^U = 0.7$, thus considering an investor seeking an exposure toward illiquid assets. For comparison purposes, we also build an alternative portfolio Π_A composed by the assets strictly laying inside the interval (π_L, π^U). To appreciate the benefits of testing the staleness level, we evaluate Π_P and Π_A over two different dimensions: first, we assess the stability of the two portfolios computing the daily portfolio turnover (Liu, 2009; Bollerslev et al., 2018),

$$\mathsf{TO}_t = \sum_{j=1}^{N} \left| w_{j,t} - w_{j,t-1} \frac{1 + r_{j,t}}{1 + \mathbf{w}_t' \mathbf{r}_t} \right|, \tag{27}$$

where $\mathbf{r}_t = (r_{1,t}, \ldots, r_{N,t})$ denotes the vector of daily returns on the N assets on day t. Our procedure takes into account the intrinsic uncertainty in the estimates of idle time and should yield a more stable asset selection over time, thus lowering the turnover of Π_P with respect to Π_A. Second, a lower turnover reduces the transaction costs thus increasing the utility of an investor. We assume the investor has quadratic utility with risk aversion γ,

$$\mathrm{U}\left(r_{p,t}, \gamma\right) = (1 + r_{p,t}) - \frac{\gamma}{2(1+\gamma)} \left(1 + r_{p,t}\right)^2,$$

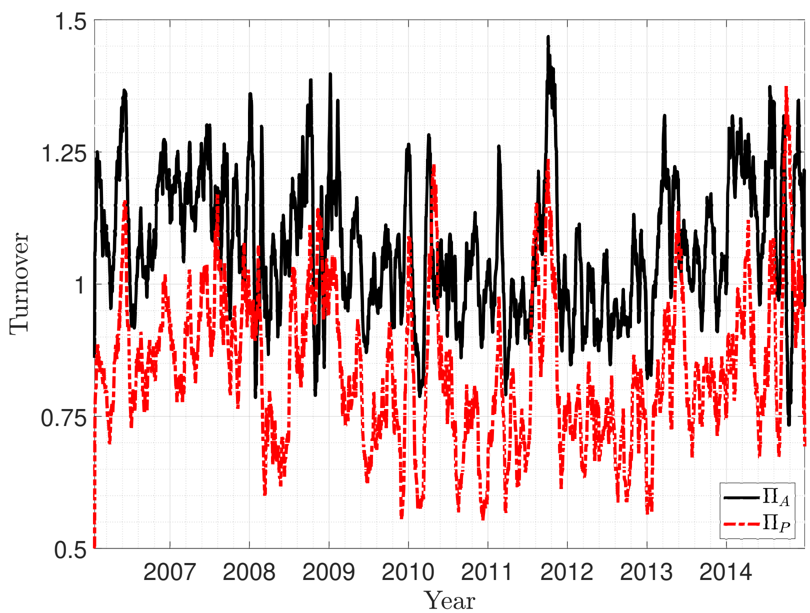
with $r_{p,t} = \mathbf{w}_t' \mathbf{r}_t - c\,\mathsf{TO}_t$ the portfolio returns adjusted by the transaction cost c. We compute the fee φ_γ an investor with risk aversion γ would be willing to pay to switch from portfolio Π_A to portfolio Π_P solving the system

$$\sum_{t=1}^{T} \mathrm{U}\left(r_{p,t}^{(\Pi_A)}, \gamma\right) = \sum_{t=1}^{T} \mathrm{U}\left(r_{p,t}^{(\Pi_P)} - \varphi_\gamma, \gamma\right).$$

To determine whether the φ_γ are significantly different from zero, we use a block bootstrap on the difference in the utilities with a block length of 12 days.

Figure 3 reports the portfolio turnover in (27) for the Π_P and the Π_A portfolio, respectively. The selection rule based on the staleness equivalence test attenuates spurious turnover, making the portfolio more stable over time and thus less expensive. To assess to what extent the lower turnover yields utility gain, we consider a coefficient of risk aversion $\gamma = \{1, 10\}$ and transaction costs $c = \{0.5\%, 1\%, 2\%\}$. These values are similar to those usually considered in the literature (Bollerslev et al., 2018). Table 6 shows that a positive and statistically significant utility gain appears regardless of the degree of transaction costs. In particular, an investor would be willing to pay up to 115 basis points per year to use the stock selection rule based on the staleness level test.

Figure 3: Portfolio turnover



Time series of portfolio turnover, as defined in Equation (27), for the Π_P portfolio (*red dotted line*), computed using the multiple staleness level test, and the Π_A portfolio (*black continuous line*), composed by the assets strictly inside the interval $(\pi_L, \pi^U) = (0.5, 0.7)$.

Table 6: Liquidity portfolio performance

c	$\gamma = 1$	$\gamma = 10$
0.5%	28.91*	28.74*
1%	57.81*	57.14*
2%	115.47*	113.05*

The table reports the economic gains in annual basis points, φ_γ, of switching from the Π_A portfolio to the Π_P portfolio for various transaction costs c and risk aversion coefficients γ. Asterisks denote φ_γ significantly different from zero at the 5% level.

5. Conclusions

Building on the econometrics framework of stale prices, we propose two statistical tests designed to study the liquidity of equity stocks. We show how to use these tests to answer two questions from the financial economic literature.

The first testing procedure allows to create groups composed by equally liquid stocks. We apply this procedure for the construction of a large dimension covariance matrix using high frequency data and prove its validity against several benchmark estimators.

The second testing procedure allows to assess whether a set of stocks has a pre-specified liquidity profile. Taking the perspective of an investor willing to bear an exposure toward illiquidity, we show that the portfolio built using our testing procedure results in a higher utility for the investor.

Acknowledgments

Davide Pirino acknowledges financial support from the project HiDEA (Advanced Econometric methods for High-frequency Data) financed by the Italian Ministry of Education, University and Research (MIUR) under the program "PRIN: progetti di ricerca di rilevante interesse nazionale - Bando 2017" Prot.

2017RSMPZZ. Alessandro Pollastri declares that the views, thoughts and opinions expressed in the text belong solely to the author, and not necessarily to the author's employer, organization, committee or other group or individual.

Appendix A. Technical Appendix

Appendix A.1. Technical assumptions from Bandi et al. (2020b)

We assume that the Bernoulli variates have limited temporal dependence, i.e.

$$\mathbb{V}\left(\tfrac{1}{n}\sum_{j=1}^{n}\prod_{q=1}^{N}\mathsf{I}_{j,n}^{(q)}\right)\xrightarrow{n\to\infty}0,$$
$$\mathbb{V}\left(\tfrac{1}{n}\sum_{j=1}^{n}\mathsf{S}_{j,n}\right)\xrightarrow{n\to\infty}0,$$

and

$$\mathbb{V}\left(\tfrac{1}{n}\sum_{j=1}^{n}\mathsf{S}_{j,n}\prod_{q=1}^{N}\mathsf{I}_{j,n}^{(q)}\right)\xrightarrow{n\to\infty}0.$$

We also assume that the maximum number of consecutive stale prices is asymptotically bounded. Let $f_j^{(q)}$ be the number of consecutive stale prices for the qth asset before instant $s_{n,j}$, with $j=1,\dots,n$, then $F_n^{(q)}=\max_j f_j^{(q)}$ is such that

$$\frac{F_n^{(q)}\log n}{n^{\alpha}}\xrightarrow{n\to\infty}0.$$

with $\alpha<1/2$.

Appendix A.2. Proof of Theorem (2.1)

Proof. We first prove the result for the test on the staleness level. From Bandi et al. (2020b), it holds that

$$\mathbb{U}_n^{(k)}=\frac{1}{n}\left(\sum_{j=1}^{n}\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}\right)+o_P\left(\frac{1}{\sqrt{n}}\right). \tag{A.1}$$

Recalling that $p_{\mathsf{T},n}^{(k)}=p_{\mathsf{S},n}+(1-p_{\mathsf{S},n})\,p_{\mathsf{I},n}^{(k)}$, we have that

$$\mathbb{E}\left[\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}\right]=p_{\mathsf{T},n}^{(k)},$$

$$\mathbb{V}\left(\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}\right)=p_{\mathsf{T},n}^{(k)}\left(1-p_{\mathsf{T},n}^{(k)}\right),$$

and

$$v_n^2=\sum_{j=1}^{n}\mathbb{V}\left(\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}-\pi\right)=n\,p_{\mathsf{T},n}^{(k)}\left(1-p_{\mathsf{T},n}^{(k)}\right).$$

As the expected value $\mathbb{E}\left[\left|\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}-p_{\mathsf{T},n}^{(k)}\right|^{2+\delta}\right]$ does not depend on j, we have that $\sum_{j=1}^{n}\mathbb{E}\left[\left|\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}-p_{\mathsf{T},n}^{(k)}\right|^{2+\delta}\right] \sim n$ as $n \to \infty$. Therefore the Lyapunov condition is satisfied, i.e.

$$\forall\delta>0 \Rightarrow \frac{1}{v_n^{2+\delta}}\sum_{j=1}^{n}\mathbb{E}\left[\left|\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}-p_{\mathsf{T},n}^{(k)}\right|^{2+\delta}\right]=O\left(\frac{1}{n^{\delta/2}}\right)\to 0,$$

the CLT applies

$$\frac{1}{v_n}\sum_{j=1}^{n}\left(\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}-p_{\mathsf{T},n}^{(k)}\right)\xrightarrow{d}\mathrm{N}(0,1),$$

and by Slutsky's theorem, we have that under $\mathcal{H}_0^L$ in (7),

$$\begin{aligned}
\left(\frac{n}{p_{\mathsf{T},n}^{(k)}\left(1-p_{\mathsf{T},n}^{(k)}\right)}\right)^{\frac{1}{2}}\ell_n^{(k)} &= \frac{\sqrt{n}}{\sqrt{p_{\mathsf{T},n}^{(k)}\left(1-p_{\mathsf{T},n}^{(k)}\right)}}\left(\mathbb{U}_n^{(k)}-\pi\right)\\
&= \frac{1}{\sqrt{np_{\mathsf{T},n}^{(k)}\left(1-p_{\mathsf{T},n}^{(k)}\right)}}\sum_{j=1}^{n}\left(\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}-\pi\right)+o_P(1)\\
&= \underbrace{\frac{1}{v_n}\sum_{j=1}^{n}\left(\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}-p_{\mathsf{T},n}^{(k)}\right)}_{\xrightarrow{d}\mathrm{N}(0,1)}+\overbrace{\frac{\sqrt{n}\left(p_{\mathsf{T},n}^{(k)}-\pi\right)}{p_{\mathsf{T},n}^{(k)}\left(1-p_{\mathsf{T},n}^{(k)}\right)}}^{\to 0}+o_P(1)\\
&\xrightarrow{d} \mathrm{N}(0,1),
\end{aligned}$$

which completes the first proof.

We now turn to the proof for the test of staleness equivalence. Note that (A.1) implies that,

$$\begin{aligned}
\mathbb{U}_n^{(q)}-\mathbb{U}_n^{(k)} &= \tfrac{1}{n}\textstyle\sum_{j=1}^{n}\left(\mathsf{S}_{j,n}+\mathsf{I}_{j,n}^{(q)}-\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(q)}-\mathsf{S}_{j,n}-\mathsf{I}_{j,n}^{(k)}+\mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}\right)+o_P(\tfrac{1}{\sqrt{n}})\\
&= \tfrac{1}{n}\textstyle\sum_{j=1}^{n}\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}-\mathsf{S}_{j,n}\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)\right)+o_P(\tfrac{1}{\sqrt{n}})\\
&= \tfrac{1}{n}\textstyle\sum_{j=1}^{n}\left(1-\mathsf{S}_{j,n}\right)\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)+o_P(\tfrac{1}{\sqrt{n}}).
\end{aligned}$$

Under $\mathcal{H}_0^E$ in (8), we have that

$$\mathbb{E}\left[(1-\mathsf{S}_{j,n})\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)\right]=0,$$

$$\mathbb{V}\left((1-\mathsf{S}_{j,n})\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)\right)=(1-p_{\mathsf{S},n})\,\left(p_{\mathsf{I},n}^{(q)}+p_{\mathsf{I},n}^{(k)}-2\,p_{\mathsf{I},n}^{(q)}\,p_{\mathsf{I},n}^{(k)}\right),$$

and follows that

$$s_n^2=\sum_{j=1}^{n}\mathbb{V}\left((1-\mathsf{S}_{j,n})\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)\right)=n(1-p_{\mathsf{S},n})\left(p_{\mathsf{I},n}^{(q)}+p_{\mathsf{I},n}^{(k)}-2p_{\mathsf{I},n}^{(q)}p_{\mathsf{I},n}^{(k)}\right).$$

Since $\left|(1-\mathsf{S}_{j,n})\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)\right|\in\{0,1\}$, then we have that for all $\delta>0$

$$\mathbb{E}\left[\left|(1-\mathsf{S}_{j,n})\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)\right|^{2+\delta}\right]=\gamma_\delta<\infty,$$

for all j, and $\sum_{j=1}^{n}\mathbb{E}\left[|\,(1-\mathsf{S}_{j,n})\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)|^{2+\delta}\right]\sim n$ as $n\to\infty$. Then the Lyapunov condition is satisfied, i.e.

$$\forall\delta>0\Rightarrow\frac{1}{s_n^{2+\delta}}\sum_{j=1}^{n}\mathbb{E}\left[|\,(1-\mathsf{S}_{j,n})\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)|^{2+\delta}\right]=O\left(\frac{1}{n^{\delta/2}}\right)\to 0.$$

and by CLT, we obtain

$$\frac{1}{s_n}\sum_{j=1}^{n}(1-\mathsf{S}_{j,n})\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)\xrightarrow{d}\mathrm{N}(0,1).$$

From Slutsky's theorem, we finally obtain

$$\begin{aligned}\left(\frac{n}{(1-p_{\mathsf{S},n})\left(p_{\mathsf{I},n}^{(q)}+p_{\mathsf{I},n}^{(k)}-2p_{\mathsf{I},n}^{(q)}p_{\mathsf{I},n}^{(k)}\right)}\right)^{\frac{1}{2}}t_n^{(q,k)} &= \frac{\sqrt{n}}{(1-p_{\mathsf{S},n})\left(p_{\mathsf{I},n}^{(q)}+p_{\mathsf{I},n}^{(k)}-2p_{\mathsf{I},n}^{(q)}p_{\mathsf{I},n}^{(k)}\right)}\left(\mathbb{U}_n^{(q)}-\mathbb{U}_n^{(k)}\right)\\ &= \frac{1}{s_n}\sum_{j=1}^{n}(1-\mathsf{S}_{j,n})\left(\mathsf{I}_{j,n}^{(q)}-\mathsf{I}_{j,n}^{(k)}\right)+o_P(1)\xrightarrow{d}\mathrm{N}(0,1),\end{aligned}$$

which concludes the proof, since all the results under the alternative hypotheses are trivial. □

Appendix A.3. Proof of Corollary (2.2)

Proof. From Bandi et al. (2020b) we have that, for all $k, q = 1, ..., N$ and $q \neq k$

$$\begin{aligned} \mathbb{U}_n^{(k)} &\xrightarrow{p} p_{\mathsf{T}}^{(k)} = p_{\mathsf{S}} + (1-p_{\mathsf{S}})\, p_{\mathsf{I}}^{(k)}, \\ \mathbb{M}_n^{(q,k)} &\xrightarrow{p} p_{\mathsf{S}} + (1-p_{\mathsf{S}})\, p_{\mathsf{I}}^{(q)} p_{\mathsf{I}}^{(k)}, \end{aligned} \tag{A.2}$$

from which

$$\begin{aligned} \mathbb{U}_n^{(k)}\left(1-\mathbb{U}_n^{(k)}\right) &\xrightarrow{p} p_{\mathsf{T}}^{(k)}\left(1-p_{\mathsf{T}}^{(k)}\right), \\ \mathbb{U}_n^{(q)} + \mathbb{U}_n^{(k)} - 2\mathbb{M}_n^{(q,k)} &\xrightarrow{p} (1-p_{\mathsf{S}})\left(p_{\mathsf{I}}^{(q)} + p_{\mathsf{I}}^{(k)} - 2p_{\mathsf{I}}^{(q)} p_{\mathsf{I}}^{(k)}\right). \end{aligned}$$

Therefore, under $\mathcal{H}_0^L$

$$\begin{aligned} \widehat{\mathbb{V}}_{\ell_n^{(k)}}^{-1/2}\, \ell_n^{(k)} &= \left(\frac{n}{\mathbb{U}_n^{(k)}\left(1-\mathbb{U}_n^{(k)}\right)}\right)^{\frac{1}{2}} \ell_n^{(k)} \\ &= \underbrace{\left(\frac{p_{\mathsf{T},n}^{(k)}\left(1-p_{\mathsf{T},n}^{(k)}\right)}{\mathbb{U}_n^{(k)}\left(1-\mathbb{U}_n^{(k)}\right)}\right)^{\frac{1}{2}}}_{\xrightarrow{p} 1} \underbrace{\left(\frac{n}{p_{\mathsf{T},n}^{(k)}\left(1-p_{\mathsf{T},n}^{(k)}\right)}\right)^{\frac{1}{2}} \ell_n^{(k)}}_{\xrightarrow{d} \mathrm{N}(0,1)} \xrightarrow{d} \mathrm{N}(0,1), \end{aligned}$$

and under $\mathcal{H}_0^E$

$$\begin{aligned} \widehat{\mathbb{V}}_{t_n^{(q,k)}}^{-1/2}\, t_n^{(q,k)} &= \left(\frac{n}{\mathbb{U}_n^{(q)} + \mathbb{U}_n^{(k)} - 2\mathbb{M}_n^{(q,k)}}\right)^{\frac{1}{2}} t_n^{(q,k)} \\ &= \underbrace{\left(\frac{(1-p_{\mathsf{S},n})\left(p_{\mathsf{I},n}^{(q)} + p_{\mathsf{I},n}^{(k)} - 2p_{\mathsf{I},n}^{(q)} p_{\mathsf{I},n}^{(k)}\right)}{\mathbb{U}_n^{(q)} + \mathbb{U}_n^{(k)} - 2\mathbb{M}_n^{(q,k)}}\right)^{\frac{1}{2}}}_{\xrightarrow{p} 1} \underbrace{\left(\frac{n}{(1-p_{\mathsf{S},n})\left(p_{\mathsf{I},n}^{(q)} + p_{\mathsf{I},n}^{(k)} - 2p_{\mathsf{I},n}^{(q)} p_{\mathsf{I},n}^{(k)}\right)}\right)^{\frac{1}{2}} t_n^{(q,k)}}_{\xrightarrow{d} \mathrm{N}(0,1)} \xrightarrow{d} \mathrm{N}(0,1). \end{aligned}$$

□

Appendix A.4. Proof of Lemma (1)

Proof. From the Cramer-Wold theorem, we know that

$$\sqrt{n}\mathcal{L}_n \xrightarrow{d} \mathbf{N}_N(0,\Sigma) \quad \text{iff} \quad \sqrt{n}\mathbf{a}'\mathcal{L}_n \xrightarrow{d} \mathrm{N}(0,\mathbf{a}'\Sigma\mathbf{a}),$$

for any (column) vector $\mathbf{a} = \left(a^{(1)}, \ldots, a^{(N)}\right)' \in \mathbb{R}^N$. Let $\mathbf{B}_{j,n}$ be defined as

$$\mathbf{B}_{j,n} \doteq \begin{bmatrix} \mathsf{S}_{j,n} + \mathsf{I}_{j,n}^{(1)} - \mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(1)} - p_{\mathsf{T},n}^{(1)} \\ \vdots \\ \mathsf{S}_{j,n} + \mathsf{I}_{j,n}^{(N)} - \mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(N)} - p_{\mathsf{T},n}^{(N)} \end{bmatrix},$$

so that, under $\mathcal{H}_0^{ML}$ in (15), it holds $\mathbb{E}\left[\mathbf{B}_{j,n}\right] = 0$ (component-wise). Then, by linearity, $\mathbb{E}\left[\mathbf{a}'\mathbf{B}_{j,n}\right] = 0$. Now consider the matrix $\mathbf{M}_n \doteq \mathbb{E}\left[\mathbf{B}_{j,n}\mathbf{B}_{j,n}'\right]$, whose generic element $\mathbf{M}_n^{(k,q)}$ is computed as

$$\begin{aligned}
\mathbf{M}_n^{(k,q)} &= \mathbb{E}\left[\left(\mathsf{S}_{j,n} + \mathsf{I}_{j,n}^{(k)} - \mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)} - p_{\mathsf{T},n}^{(k)}\right)\left(\mathsf{S}_{j,n} + \mathsf{I}_{j,n}^{(q)} - \mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(q)} - p_{\mathsf{T},n}^{(q)}\right)\right] \quad \forall k, q \in \{1, \ldots, N\} \\
&= \mathbb{E}\left[\left(\mathsf{S}_{j,n} + \mathsf{I}_{j,n}^{(k)} - \mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(k)}\right)\left(\mathsf{S}_{j,n} + \mathsf{I}_{j,n}^{(q)} - \mathsf{S}_{j,n}\mathsf{I}_{j,n}^{(q)}\right)\right] - p_{\mathsf{T},n}^{(k)}p_{\mathsf{T},n}^{(q)} \quad \forall k, q \in \{1, \ldots, N\} \\
&= \begin{cases} p_{\mathsf{T},n}^{(k)}\left(1 - p_{\mathsf{T},n}^{(k)}\right) & k = q \\ p_{\mathsf{S},n} + (1 - p_{\mathsf{S},n})p_{\mathsf{I},n}^{(q)}p_{\mathsf{I},n}^{(k)} - p_{\mathsf{T},n}^{(k)}p_{\mathsf{T},n}^{(q)} & k \neq q \end{cases}
\end{aligned}$$

Since $\mathbb{E}\left[\mathbf{B}_{j,n}\right] = 0$ we have

$$\mathbb{V}\left[\mathbf{a}'\mathbf{B}_{j,n}\right] = \mathbf{a}'\mathbb{E}\left[\mathbf{B}_{j,n}\mathbf{B}_{j,n}'\right]\mathbf{a} = \mathbf{a}'\mathbf{M}_n\mathbf{a},$$

and follows that

$$q_n^2 = \sum_{j=1}^{n} \mathbb{V}\left[\mathbf{a}'\mathbf{B}_{j,n}\right] = n\, \mathbf{a}'\mathbf{M}_n\mathbf{a}.$$

As $\mathbb{E}\left[\left|\mathbf{a}'\mathbf{B}_{j,n}\right|^{2+\delta}\right]$ does not depend on j, then $\sum_{j=1}^{n} \mathbb{E}\left[\left|\mathbf{a}'\mathbf{B}_{j,n}\right|^{2+\delta}\right] \sim n$, as $n \to \infty$. Hence, for all $\delta > 0$

$$\tfrac{1}{q_n^{2+\delta}} \textstyle\sum_{j=1}^{n} \mathbb{E}\left[\left|\mathbf{a}'\mathbf{B}_{j,n}\right|^{2+\delta}\right] = O\left(\tfrac{1}{n^{\delta/2}}\right) \to 0,$$

and the Lyapunov's condition is thus satisfied. Applying the CLT, we obtain

$$\frac{1}{q_n} \sum_{j=1}^{n} \mathbf{a}'\mathbf{B}_{j,n} \xrightarrow{d} \mathrm{N}(0, 1).$$

Under $\mathcal{H}_0^{ML}$ in (15), we have that $\mathbf{M}_n \to \Sigma$, where Σ is defined in Equation (19), and by Slutsky's theorem

$$\begin{aligned}
\sqrt{n}\mathbf{a}'\mathcal{L}_n &= \sqrt{n}\left[a^{(1)}\left(\mathbb{U}_n^{(1)}-\pi\right)+\cdots+a^{(N)}\left(\mathbb{U}_n^{(N)}-\pi\right)\right] \\
&\overset{\text{Using (A.1)}}{=} \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\mathbf{a}'\mathbf{B}_{j,n}+\sqrt{n}\left[a^{(1)}\left(p_{\mathsf{T},n}^{(1)}-\pi\right)+\cdots+a^{(N)}\left(p_{\mathsf{T},n}^{(N)}-\pi\right)\right]+o_P(1) \\
&= \frac{\sqrt{\mathbf{a}'\mathbf{M}_n\mathbf{a}}}{\sqrt{n\mathbf{a}'\mathbf{M}_n\mathbf{a}}}\sum_{j=1}^{n}\mathbf{a}'\mathbf{B}_{j,n}+\sqrt{n}\left[a^{(1)}\left(p_{\mathsf{T},n}^{(1)}-\pi\right)+\cdots+a^{(N)}\left(p_{\mathsf{T},n}^{(N)}-\pi\right)\right]+o_P(1) \\
&= \underbrace{\sqrt{\mathbf{a}'\mathbf{M}_n\mathbf{a}}}_{\longrightarrow\sqrt{\mathbf{a}'\Sigma\,\mathbf{a}}}\underbrace{\frac{1}{q_n}\sum_{j=1}^{n}\mathbf{a}'\mathbf{B}_{j,n}}_{\xrightarrow{d}\mathrm{N}(0,1)}+\sqrt{n}\underbrace{\left[a^{(1)}\left(p_{\mathsf{T},n}^{(1)}-\pi\right)+\cdots+a^{(N)}\left(p_{\mathsf{T},n}^{(N)}-\pi\right)\right]}_{\longrightarrow 0}+o_P(1) \\
&\xrightarrow{d} N\left(0,\mathbf{a}'\Sigma\mathbf{a}\right),
\end{aligned}$$

which completes the proof. Analogous steps can be used to show that under $\mathcal{H}_0^{ME}$ in (16), $\sqrt{n}\mathcal{T}_n \xrightarrow{d} \mathbf{N}_{\frac{N(N-1)}{2}}(0,\Xi)$.

□

References

Arellano-Valle, R. B., & Genton, M. G. (2008). On the exact distribution of the maximum of absolutely continuous dependent random variables. *Statistics & Probability Letters*, *78*, 27–35.

Bandi, F., Kolokolov, A., Pirino, D., & Renó, R. (2020a). Zeros. *Management Science*, *66*, 1–14.

Bandi, F., Pirino, D., & Renó, R. (2017). Excess idle time. *Econometrica*, *85*, 1793–1846.

Bandi, F., Pirino, D., & Renó, R. (2020b). Systematic staleness. *SSRN*, .

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

Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2018). Modeling and forecasting (un) reliable realized covariances for more reliable financial decisions. *Journal of Econometrics*, *207*, 71–91.

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events: for insurance and finance* volume 33. Springer Science & Business Media.

Engle, R. F., & Colacito, R. (2006). Testing and valuing dynamic correlations for asset allocation. *Journal of Business and Economics Statistics*, *24*, 238–253.

Fleming, J., Kirby, C., & Ostdiek, B. (2001). The economic value of volatility timing. *The Journal of Finance*, *56*, 329–352.

Fleming, J., Kirby, C., & Ostdiek, B. (2003). The economic value of volatility timing using "realized" volatility. *Journal of Financial Economics*, *67*, 473–509.

Golosnoy, V., Schmid, W., Seifert, M. I., & Lazariv, T. (2020). Statistical inferences for realized portfolio weights. *Econometrics and Statistics*, *14*, 49–62. URL: https://www.sciencedirect.com/science/article/pii/S2452306218300534. doi:https://doi.org/10.1016/j.ecosta.2018.08.003.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*, 453–497.

Hautsch, N., Kyj, L. M., & Malec, P. (2015). Do high-frequency data improve high-dimensional portfolio allocations? *Journal of Applied Econometrics*, *30*, 263–290.

Hautsch, N., Kyj, L. M., & Oomen, R. (2012). A blocking and regularization approach to high-dimensional realized covariance estimation. *Journal of Applied Econometrics*, *27*, 625–645.

Ibbotson, R. G., Chen, Z., Kim, D. Y.-J., & Hu, W. Y. (2013). Liquidity as an investment style. *Financial Analysts Journal*, *69*, 30–44.

Jansen, R., & Van Dijk, R. (2002). Optimal benchmark tracking with small portfolios. *The Journal of Portfolio Management*, *28*, 33–39.

Lesmond, D. A., Ogden, J. P., & Trzcinka, C. A. (1999). A new estimate of transaction costs. *The Review of Financial Studies*, *12*, 1113–1141.

Liu, Q. (2009). On portfolio optimization: How and when do we benefit from high-frequency data? *Journal of Applied Econometrics*, *24*, 560–582.

Romano, J. P., Shaikh, A. M., & Wolf, M. (2008). Formalized data snooping based on generalized error rates. *Econometric Theory*, *24*, 404–447.

Sass, J., & Thös, A.-K. (2021). Risk reduction and portfolio optimization using clustering methods. *Econometrics and Statistics*, . URL: `https://www.sciencedirect.com/science/article/pii/S2452306221001416`. doi:`https://doi.org/10.1016/j.ecosta.2021.11.010`.