Zeros

Federico M. Bandi* Davide Pirino† Roberto Renò‡

PRELIMINARY DRAFT

Abstract

Asset prices can be stale. We define price "staleness" as lack of price adjustments yielding zero returns (i.e., zeros). The term "idleness" (resp. "near idleness") is, instead, used to define staleness when trading activity is absent (resp. close to absent). We show that zeros are a genuine economic phenomenon linked to trading volumes and liquidity. Zeros are, in general, not the result of institutional features, like price discreteness. Spells of idleness or near idleness are stylized facts suggestive of a key, omitted market friction in the modeling of asset prices.

Keywords: Volume, liquidity, continuous-time semimartingales.

*Johns Hopkins University, Carey Business School, 100 International Drive, Baltimore MD 20202, USA and Edhec-Risk Institute, France. E-mail: fbandi1@jhu.edu.

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

‡Università di Verona, Dipartimento di Scienze Economiche, Via Cantarane 24, 37129, Verona, Italy. E-mail: roberto.reno@univr.it.

1 Introduction

The ubiquitous semi-martingale model for asset prices in continuous time does not appear to be rich enough. Figure 1 displays examples of intraday price paths, sampled at 30−second intervals, for stocks traded on the New York Stock Exchange (NYSE). In the figure, the presence of trading volume is represented by blue circles whose width is proportional to dollar value (see the legend in Panel E). Absence of volume is, instead, represented by red crosses.

Panel A shows the price dynamics of a liquid stock (Exxon Mobil, XOM) with large volumes. The graph is visually compatible with the erratic behavior of a semi-martingale in continuous time, one which is locally driven by Brownian motion and, infrequently, by jump-like discontinuities. *Staleness*, defined as the frequency of zero returns, is only 3.7%. *Idleness*, defined as the frequency of time intervals without trading volume is 0.04%. (Of course, idleness is lower than staleness by construction, since absence of trading implies staleness.) Panel A illustrates an ideal situation, supporting semi-martingale modeling, one which is - however - far from typical. Specifically, Panel A displays the day in our data for which XOM has the lowest level of staleness in the roughly 10-year sample we analyze. (Details on the data sample are provided below.)

Panel B represents a more typical situation for another very liquid stock (Bank of America: BAC). The graph now looks rather different. The price process is considerably stickier. Staleness is 75.1%, idleness is 44.1%. The number associated with idleness (44.1%) raises the question of why market participants were, for nearly half of the day, electing not to trade this liquid stock. We will return to this issue. The difference between staleness and idleness (31%) is suggestive of the fact that one can often observe zeros when volumes are low, but non-zero. In Panel B, in fact, we observe relatively larger volume being associated with price changes and relatively smaller volume being associated with zeros. What could lead to zeros in the presence of small, but non-zero, volume? We identify and discuss two channels: easy absorption of limited volume without price impacts (near idleness) and/or price discreteness. While the first channel is economically relevant, the second one may not be. At the NYSE, since 2001, traded prices change by multiples of 1 cent. Zeros observed in conjunction with non-zero trading volume may, therefore, be due to this institutional feature.

This observation is illustrated in Panel C, where the dynamics of Citigroup (C) are affected by sizable staleness (88.7%) but almost no idleness (2.1%). We note that the price range of 8 cents for that day is small as compared to the minimum allowed price change of 1 cent. In effect, a large number of zeros (leading to limited price variability over the day) occur in the presence of non-zero volume. While price discreteness is the reason for this finding, should one view the effect as being spurious (i.e., economically uninformative)? We argue that price discreteness in the presence of limited volume (like in Panel C) cannot be viewed as spuriously leading to zeros. It only translates into rounded up prices which may have moved somewhat (but by less than a cent) if they had been allowed to take on a continuum of values. In Panel C, when traded volumes increase just before 4 pm, the impact of price discreteness disappears. So does the number of zeros in that time frame, due to bid/ask bounce effects and prices reverting on the discretization grid. In essence, then, for price discreteness to represent an institutional feature leading to spurious zeros *uninformative* about the nature of the trading mechanism, it would need to also be associated with relatively large volumes. If this were the case, then zeros would not just be the result of low volumes and limited liquidity. They would also be the result of high liquidity, and easily-absorbed high volumes, yielding low price impacts and (rounded up, due to price discreteness) stale traded prices. The latter scenario, however, is a rare occurrence in the data.

Panels D and E feature two stocks (VeriFone Systems: PAY and Kansas City Southern: KSU) for which the intraday price range is large with respect to 1 cent. Idleness is now strongly in the data. In the case of PAY, one could observe zeros with limited volumes (in 29.7% of the cases). As for KSU, virtually all zeros are associated with zero volumes. Thus, staleness and idleness almost coincide for this stock and this specific day in the sample. In light of the larger difference between staleness and idleness, price discreteness is more likely to affect PAY than KSU during the reported days. However, given the limited volumes associated with zero returns for PAY, price discreteness simply replaces continuous prices (and non-zero, but likely small, returns) with rounded up prices (and zero returns). Hence, in agreement with our previous observations, price discreteness does not appear to run counter to the positive relation between zeros and low volumes.

Panel F shows the price dynamics of Fluor (FLR), a stock whose price is extremely stale during the first part of the day and rather erratic (in association with larger volumes) during the second part of the day. We note that a jump around 2 pm sparks an increase in trading activity for this stock.

Figure 1 shows that idleness and near idleness should be an integral part of a realistic data generating process for asset prices in continuous time *before* institutional features (like rounding) are accounted for. The figure also raises questions. How do idleness and staleness change dynamically, intra-daily for example? What is the impact of sampling frequency on them? What are their economic determinants? We turn to some answers.

2 An upper bound on the impact of price discreteness

A portion of staleness is due to rounding. As is well-known, allowing for rounding when modeling *observed* prices (given a continuous-time model for latent fundamentals) is a challenging issue (e.g., Delattre and Jacod (1997)). In fact, the probability of a zero return, conditional on the past history of prices, is a strongly non-Markovian problem, even under simple assumptions on the price dynamics.

Having made this point, we may provide an expression for the conditional (given time t information) probability of observing a zero return due to rounding using an uninformative uniform prior for the location of latent fundamentals within the discretization grid. Express the latent price P_t as the discretized analogue (over an interval of length Δ) of the local martingale

$$P_t = \int_0^t \sigma_s P_s dW_s, \tag{1}$$

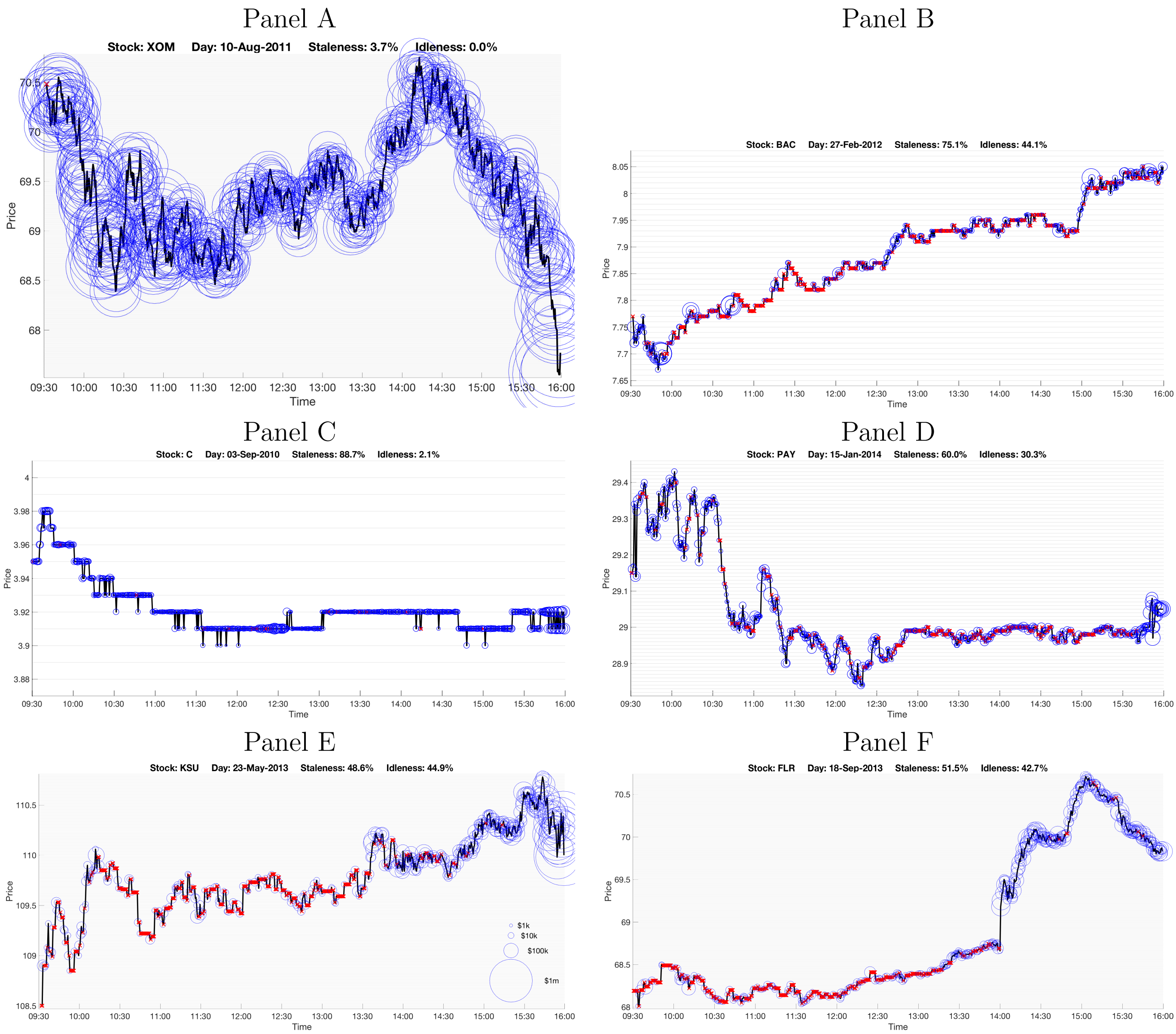

Figure 1: Six examples of intraday stock price dynamics on a 30-second sampling grid. In each panel, the circles represent dollar volume traded in each 30-second interval. The sizes of the circles represent dollar values, as indicated in Panel E. Absence of volume in a 30-second interval is signaled by a red cross. Values that prices can take (as multiples of 1 cent) are represented on the horizontal grid.

where W_t is Brownian motion. We are neglecting the drift term (which is reasonable, over small intervals, if the drift is uniformly bounded) and price discontinuities (which is also reasonable, if these discontinuities are rare). Eq. (1) implies that the logarithmic price process is locally Gaussian with instantaneous volatility σ_t.

We now assume that the traded price $\overline{P}_t$ can be only observed as a multiple of a fixed quantity d (e.g., for NYSE-listed stocks d is one cent) and that $\overline{P}_t$ is obtained by rounding P_t to the nearest value in the set $\{0, d, 2d, \ldots, kd, \ldots\}$. The conditional probability that $\overline{P}_{t+\Delta} - \overline{P}_t = 0$, i.e., the probability of observing a zero return because of rounding, is

$$\mathbb{P}\left(\overline{P}_t - d/2 < P_{t+\Delta} < \overline{P}_t + d/2 | P_t = \overline{P}_t + x\right) = \int_{-(d/2+x)}^{d/2-x} \frac{1}{\sqrt{2\pi\sigma_t^2 P_t^2 \Delta}} e^{-\frac{z^2}{2\sigma_t^2 P_t^2 \Delta}} dz,$$

where $-d/2 < x < d/2$. Assume, now, that $x = P_t - \overline{P}_t$ is uniformly distributed over $[-d/2, d/2]$. In other words, when we observe the rounded price $\overline{P}_t$, we do not have any information as to the location of P_t in the interval $[\overline{P}_t - d/2, \overline{P}_t + d/2]$. Given this assumption, integrating x out, the conditional probability of observing a zero return due to rounding, denoted by $p_t^{\varnothing,R}$, is

$$p_t^{\varnothing,R} = \frac{1}{d}\int_{-d/2}^{d/2}\int_{-(d/2+x)}^{d/2-x} \frac{1}{\sqrt{2\pi\sigma_t^2 P_t^2 \Delta}} e^{-\frac{z^2}{2\sigma_t^2 P_t^2 \Delta}} dz dx = \operatorname{erf}\left(\frac{d}{\sqrt{2\sigma_t^2 P_t^2 \Delta}}\right) + \sqrt{\frac{2}{\pi}}\left(e^{-\frac{d^2}{2\sigma_t^2 P_t^2 \Delta}} - 1\right)\frac{\sigma_t P_t \sqrt{\Delta}}{d},$$

where the symbol $\operatorname{erf}(x)$ defines the Gaussian error function. Using Eq. (2), a stock with, for example, price $P = 10$, daily volatility of 1% (one day = 6.5 hours for NYSE stocks) and discretization $d = 0.01$ will have a zero return due to price discreteness with probability 4% over the next day, 10.12% over the next hour, 60.52% over the next minute, and 94.78% over the next second.

We note that the lower volatility (σ_t), the larger the probability of price discreteness ($p_t^{\varnothing,R}$). Low volatility is, however, generally associated with low volumes (c.f., Tauchen and Pitts (1983) and the references therein). Thus, rounded up traded prices (and, as a consequence, zero returns) are hardly in contradiction with slow trading. Once more, the effect of price discreteness would be spurious if it were to occur along with large volumes, something which is typically not found in the data. In this sense, Eq. (2), which is agnostic about volumes, may be viewed as a (possibly very generous) upper bound on the spurious impact of price discreteness on genuine zeros. The inclusion of bid/ask bounce effects in the data generating process would further reduce $p_t^{\varnothing,R}$.

3 *Staleness* and *idleness* in the data

Our data consists of all trades of 244 NYSE-listed stocks, recorded from 9:30am to 4pm, from 2006 to 2015. We employ the 244 stocks with the largest average traded volume during the period. In this sense, given the impact of (low) volume on staleness, the empirical findings in this study should be interpreted as conservative.

For each stock k (with $k = 1, \ldots, N$) and for each day t_k (with $t_k = 1, \ldots, T_k$), we sample prices $\overline{P}_{0,t_k} \ldots, \overline{P}_{n,t_k}$ on an evenly spaced grid and compute intraday returns as $r_{i,t_k} = \ln \overline{P}_{i,t_k} - \ln \overline{P}_{i-1,t_k}$, where $i = 1, \ldots, n$. A return is *stale* if $r_{i,t_k} = 0$. It is *idle* if $V_{i,t_k} = 0$, where V_{i,t_k} is the traded dollar volume over the time interval $[(i-1)/n, i/n]$. (We use daily units for time.)

We note that the implementation of (2) requires estimation of a time-varying spot variance. We identify the latter as follows. First, we compute daily realized variances by summing the squares of

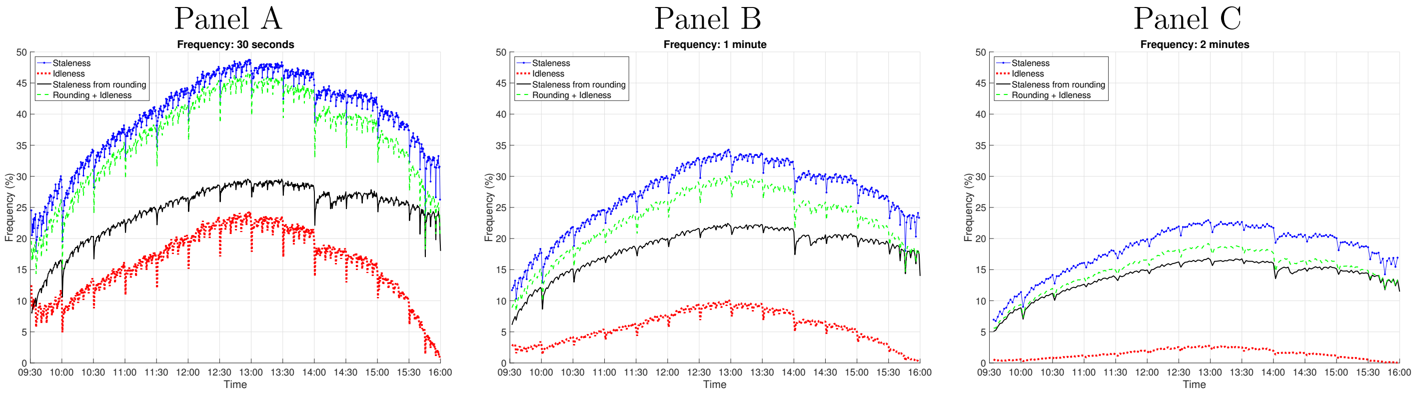

Figure 2: We report average intraday staleness (blue lines), average intraday idleness (red lines), estimated intraday staleness from rounding (black lines) and the sum of average staleness and estimated staleness from rounding (green lines). Prices are sampled every 30 seconds (Panel A), every 1 minute (Panel B) and every 2 minutes. *Staleness* is the frequency of zero returns. *Idleness* is the frequency of zero returns without volume. Estimated staleness from rounding is computed as in Eq. (2).

the intraday logarithmic returns

$$RV_{t_k} = \sum_{i=1}^{n} r_{i,t_k}^2.$$

Second, we define $\widehat{\sigma}_{i,t_k}^2 = w_i RV_{t_k}$ for each intra-daily return on day t_k, where w_i is a scaling factor designed to account for intra-daily effects:

$$w_i = \frac{1}{T_k} \sum_{t_k=1}^{T_k} \frac{r_{i,t_k}^2}{RV_{t_k}}.$$

Figure 2 reports average intraday staleness (blue lines), average intraday idleness (red lines) and the *estimated* (using Eq. (2)) upper bound on staleness due to price discreteness (black lines) at three frequencies: 30 seconds (Panel A), 1 minute (Panel B) and 2 minutes (Panel C). Both staleness and idleness display a pronounced inverse-U shape mirroring lower volatility and lower trading activity around lunch time. Both quantities are also smaller (on average) every five minutes and, in particular, every half hour, an empirical pattern attributed by some to human biases in algorithmic trading (Broussard and Nikiforov (2013)). They are, moreover, lower around 10:00am, reflecting the release of some key macroeconomic announcements at this time. As expected, average staleness and idleness decrease with the sampling frequency. At their peak (lunch time) staleness and idleness are about 50% and 25% with 30-second sampling, 35% and 10% with 1-minute sampling and 22% and 3% with 2-minute sampling. The higher the sampling frequency, the more the adjustment in Eq. (2) fills the gap between staleness and genuine idleness (zero returns due to zero volumes). When sampling every 30 seconds, for example, the blue line (total staleness) is very close to the green line (the sum of idleness and the estimated upper bound on staleness due to price discreteness). We

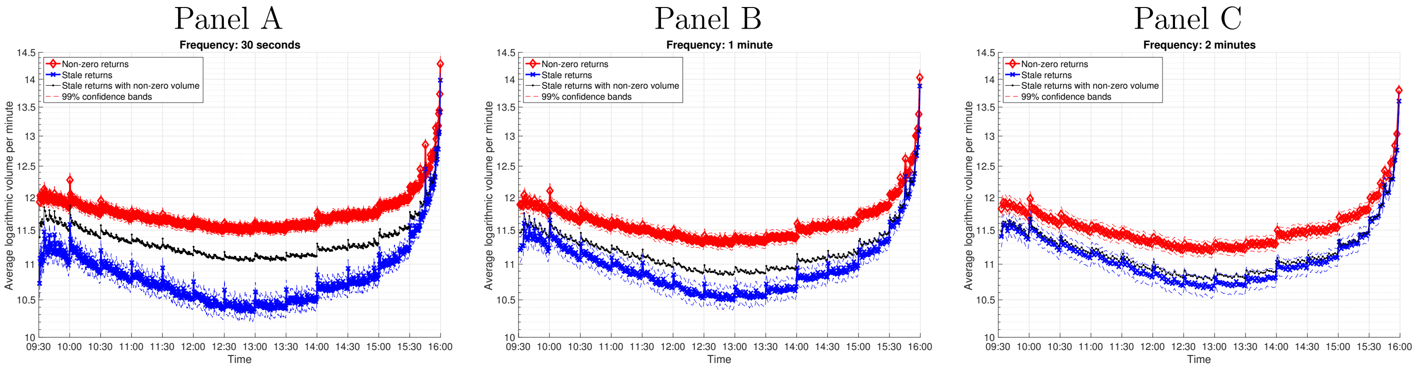

Figure 3: We report average intraday volume associated with non-zero returns (red line) with 99% confidence bands, average intraday volume associated with zero returns (blue line) with 99% confidence bands and average intraday volume associated with zero returns with non-zero volume. We consider three sampling frequencies: 30 seconds, 1 minute and 2 minutes.

interpret this result as suggesting that, at lower frequencies, there is a component of staleness which is not spurious (i.e., due to rounding) and is, also, not due to zero volumes, i.e. idleness. Said differently, we may observe genuine staleness in the presence of low (but strictly non-zero) returns. This additional (third) component of staleness, a sort of *near idleness*, is relatively more important the lower the sampling frequency.

4 The economics of staleness

What generates staleness and idleness? As pointed out, rounding leads to some staleness. However, at all reported frequencies, staleness is significantly higher than its spurious component due to rounding (even when employing an upper bound on the latter). Bandi et al. (2017) provide a possible economic rationale for genuine staleness which hinges on micro-structural theories of price formation with transaction costs and asymmetries in information (c.f. Hasbrouck and Ho (1987), Kyle (1985) and Glosten and Milgrom (1985)). The logic in Bandi et al. (2017) is as follows: the informed traders, by definition, "know" the fundamental values. These values are, in fact, expectations of future cash flows given their information sets. The trading decisions of these traders are, therefore, based on the *relative* difference between, on the one hand, the gap between midpoints of bid/ask spreads and fundamental values and, on the other hand, quoted (and perceived) execution costs, inclusive of (half) bid/ask spreads. Specifically, these traders will execute trades only when the trades guarantee a profit, net of execution costs. Should execution costs be excessively large, they will be reluctant to trade. Uninformed traders may also respond to the magnitude of execution costs and trade less when these costs are perceived to be too large. Because fundamental values are however unknown to them, their trading decisions will, in general, be based on the *absolute* magnitude of the cost of transacting.

This logic provides economic support for absence of trading (particularly at high frequencies) or very limited trading (particularly at low frequencies) as a source of zero returns. Since execution costs capture at least one dimension of liquidity ("tightness", rather than "depth" or "resiliency"), it also provides economic support for a positive relation between illiquidity and zero returns with volumes representing the mediating variable. We will return to these observations in a later section of the paper.

Once more, with 30-second sampling, the sum of idleness and the spurious component of staleness explains almost all of the staleness in the data. With sparser sampling, however, near idleness tends to replace pure idleness. A careful look at volumes will, now, be economically revealing. Figure 3 displays the average intraday volume (with 99% confidence bands) associated with zero returns (blue lines) and non-zero returns (red lines) for the same three frequencies considered in Figure 2: 30 seconds, 1 minute and 2 minutes. We note that intraday average volume has a U-shaped pattern which nicely reflects the inverse U-shaped pattern in staleness (c.f., Figure 2). Upward spikes in volume every 5 minutes (and, in particular, every half hour) also mirror downward spikes in staleness at the same times.

Importantly, the average volume associated with stale returns is (statistically) significantly lower than the average volume associated with non-zero returns at all frequencies and, especially, at high frequencies. Because stale returns are often idle, that is with zero volume, we also report the average intraday volume of stale returns with non-zero volume (black lines). The black line is bound to capture some price discreteness (since non-zero volume can be associated with stale prices, when price discreteness plays a role) but also near idleness (since absence of trading or limited trading leads to staleness, consistent with the logic in Bandi et al. (2017)). Not surprisingly, as we move to lower frequencies, a larger proportion of stale returns is characterized by strictly positive (i.e., non-zero) volumes. Yet, the separation between the lower (possibly non-zero, particularly at lower frequencies) volumes associated with stale returns and the higher volumes associated with non-stale returns is evident.

5 An alternative data generating process

Idleness and near idleness are stylized empirical features of recorded asset prices. Modeling prices as continuous semi-martingales (before rounding) will therefore not capture effects which are strongly in the data. Below, we provide an alternative modeling approach.

We will still assume that the *latent* price (P) follows a local martingale as in Eq. (1). Adding a finite variation drift component and/or a jump component can be easily done.

Next, we partition the sampling span $[0,T]$ into n intervals of length Δ_n. Naturally, Δ_n denotes sampling frequency. Idleness can now be modeled, as in Bandi et al. (2017), using a triangular array of Bernoulli variates $\{\mathbb{B}_{i,n}\}_{i=1,\dots,n}$ satisfying, for all $i = 1, \dots, n$,

$$\mathbb{P}(\mathbb{B}_{i,n} = 1) = \mathbb{E}(\mathbb{B}_{i,n}) = p_n \tag{2}$$

and, as $n \to \infty$,

$$\frac{1}{T}\sum_{i=1}^{n} \Delta_n \mathbb{B}_{i,n} \xrightarrow{p} p_\infty. \tag{3}$$

The *observed* price process (before rounding) on the (sampling) partition follows:

$$\widetilde{P}_{i\Delta_n} = P_{i\Delta_n}(1 - \mathbb{B}_{i,n}) + \widetilde{P}_{(i-1)\Delta_n}\mathbb{B}_{i,n}. \tag{4}$$

Thus, the observed price is the fundamental (*latent*) price if $\mathbb{B}_{i,n} = 0$ (i.e., in the presence of trading). Absent trading, i.e., if $\mathbb{B}_{i,n} = 1$, the previous price is repeated and prices are stale.[1] Of course, a more nuanced interpretation consistent with empirical evidence would set $\mathbb{B}_{i,n} = 1$ not only when trading is absent (idleness) but, also, when trading is slow (near idleness). We abide by this interpretation in what follows.

A key implication of the model is that the latent price continues to be diffusive (with some discontinuities, if jumps are allowed) *even* in the presence of staleness. A consequence of this implication, and an interesting prediction of the model, is that the variance of intraday returns should be larger after staleness than after non-zero returns. In fact, the longer prices have been stale, the larger variance should be. More precisely, ignoring jumps, variance should be an almost linear function of the extent of past staleness. Figure 4 verifies this prediction with data. For each stock and each day, we compute the relative variance of the intraday returns with respect to the overall variance for the day. We then calculate averages conditional on the number of stale returns observed before the considered return. As can be seen in the figure, the unconditional intraday (relative) variance is larger than that associated with returns not preceded by staleness. The reason is that, after spells of staleness, variance is higher, as predicted by the model. In addition, as we condition on the number of previous stale returns, variance increases in an almost linear fashion.

6 *Excess* staleness

Define *excess staleness* as total staleness net of spurious contaminations due to price discreteness. Under the assumption that only price discreteness (combined with high liquidity) can give rise to staleness in the presence of large volumes, we interpret excess staleness as genuine staleness due to zero or low volumes (i.e., the sum of idleness and near idleness). We turn to a numerical assessment.

We begin with some definitions. TS is total staleness, that is

$$\text{TS}_{j,j+1} = \{\omega \in \Omega | \widetilde{P}_{(j+1)\Delta_n} = \widetilde{P}_{j\Delta_n}\}.$$

[1] Returning to economic logic, one could argue that, in the presence of trading, fundamental values would be revealed to uninformed, or partially-informed, traders since, given Eq. (4), traded prices coincide with fundamental values (ignoring rounding). An easy way to prevent full observability of fundamental values when trading takes place is to assume $\widetilde{p}_{i\Delta_n} = p^*_{i\Delta_n}(1 - \mathbb{B}_{i,n}) + \widetilde{p}_{(i-1)\Delta_n}\mathbb{B}_{i,n}$, where $p^* = p + \epsilon$, with ϵ representing a market microstructure contamination, possibly induced by bid/ask bounce effects (see, e.g., Bandi et al. (2017)).

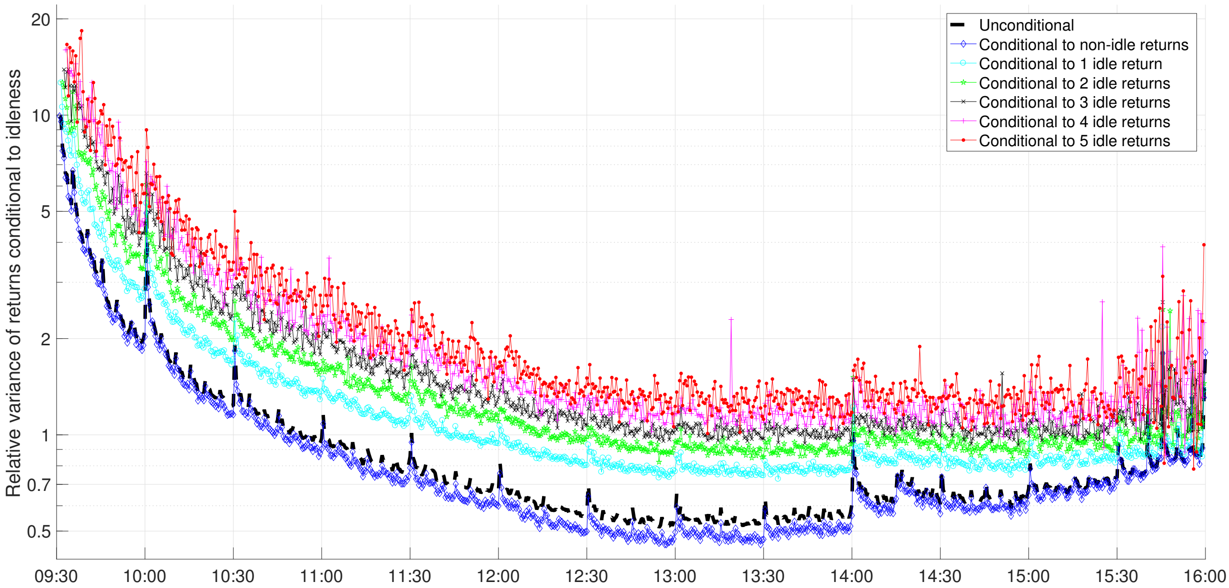

Figure 4: We report the relative variance of intraday returns with respect to the overall variance. We do so unconditionally and after conditioning on different numbers of consecutive zero returns.

R defines rounding, namely

$$\mathrm{R}_{j,j+1} = \{\omega \in \Omega | \widetilde{P}_{(j+1)\Delta_n} = \widetilde{P}_{j\Delta_n}\} \cap \{\omega \in \Omega | \mathbb{B}_{j+1,n} = 0\}.$$

ES denotes excess staleness, i.e.,

$$\mathrm{ES}_{j,j+1} = \{\omega \in \Omega | \mathbb{B}_{j+1,n} = 1\}.$$

$\mathrm{ES}^c_{j,j+1}$ is, therefore, the event "not excess staleness", i.e.,

$$\mathrm{ES}^c_{j,j+1} = \Omega \backslash \mathrm{ES}_{j,j+1} = \{\omega \in \Omega | \mathbb{B}_{j+1,n} = 0\}.$$

Given these definitions, the probability of total staleness can be expressed as follows:

$$\begin{aligned}
\mathbb{P}(\mathrm{TS}_{j,j+1}) &= \mathbb{P}(\mathrm{ES}_{j,j+1} \cup \mathrm{R}_{j,j+1}) \\
&= \mathbb{P}(\mathrm{ES}_{j,j+1}) + \mathbb{P}(\mathrm{R}_{j,j+1}) - \mathbb{P}(\mathrm{ES}_{j,j+1} \cap \mathrm{R}_{j,j+1}) \\
&= \mathbb{P}(\mathrm{ES}_{j,j+1}) + \mathbb{P}(\mathrm{R}_{j,j+1} \cap \mathrm{ES}_{j,j+1}) + \mathbb{P}(\mathrm{R}_{j,j+1} \cap \mathrm{ES}^c_{j,j+1}) - \mathbb{P}(\mathrm{ES}_{j,j+1} \cap \mathrm{R}_{j,j+1}) \\
&= \mathbb{P}(\mathrm{ES}_{j,j+1}) + \mathbb{P}(\mathrm{R}_{j,j+1} \cap \mathrm{ES}^c_{j,j+1}) \\
&= \mathbb{P}(\mathrm{ES}_{j,j+1}) + \mathbb{P}(\mathrm{R}_{j,j+1} | \mathrm{ES}^c_{j,j+1})(1 - \mathbb{P}(\mathrm{ES}_{j,j+1})) \\
&= p_n + \mathbb{P}(\mathrm{R}_{j,j+1} | \mathrm{ES}^c_{j,j+1})(1 - p_n).
\end{aligned}$$

Next, we apply rounding to the new data generating process in Eq. (4) rather than to the latent

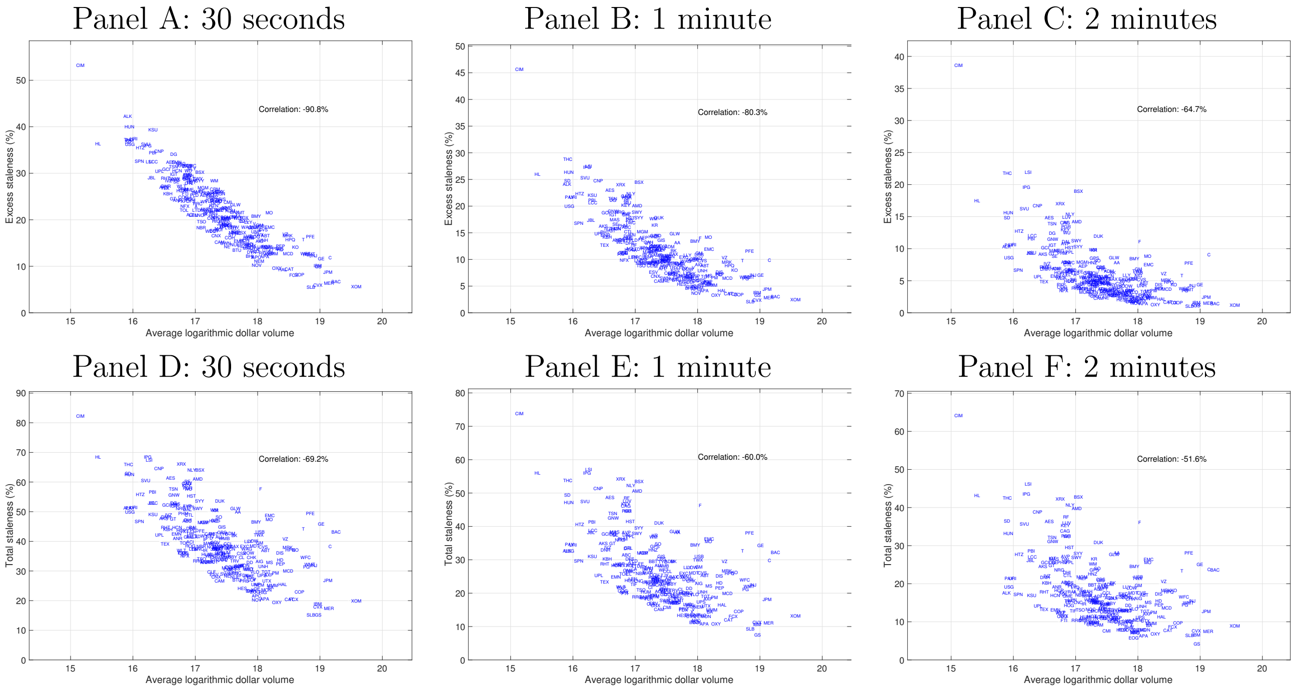

Figure 5: We plot excess staleness, i.e., the difference between average total staleness and estimated staleness from rounding, versus average dollar volume for all stocks in our sample. Total staleness is the frequency of zero returns. Estimated staleness from rounding is computed as in Eq. (2). We consider three frequencies, 30 seconds (Panel A), 1 minute (Panel B) and 2 minutes (Panel C). We also plot average total staleness versus average dollar volume for the same three frequencies.

process in Eq. (1). For notational convenience, write $\overline{P}_{j\Delta_n} - d/2 = \overline{P}^-_{j\Delta_n}$ and $\overline{P}_{j\Delta_n} + d/2 = \overline{P}^+_{j\Delta_n}$. Given $-d/2 < x < d/2$:

$$\mathbb{P}(\mathrm{R}_{j,j+1}) = \frac{1}{d}\int_{-d/2}^{d/2} \mathbb{P}\left(\overline{P}^-_{j\Delta_n} < \widetilde{P}_{(j+1)\Delta_n} < \overline{P}^+_{j\Delta_n} | \widetilde{P}_{j\Delta_n} = \overline{P}_{j\Delta_n} + x\right) dx.$$

What this means is that, conditional on not having excess staleness between $j\Delta_n$ and $(j+1)\Delta_n$, the true price coincides with the latent price and

$$\mathbb{P}(\mathrm{R}_{j,j+1}|\mathrm{ES}^c_{j,j+1}) = \frac{1}{d}\int_{-d/2}^{d/2} \mathbb{P}\left(\overline{P}^-_{j\Delta_n} < P_{(j+1)\Delta_n} < \overline{P}^+_{j\Delta_n} | \widetilde{P}_{j\Delta_n} = \overline{P}_{j\Delta_n} + x\right) dx.$$

However, we are conditioning on $\widetilde{P}_{j\Delta_n}$ not on $P_{j\Delta_n}$. This said, it holds that

$$\widetilde{P}_{j\Delta_n} = P_{(j-\mathbb{K}_{j,n})\Delta_n},$$

where, for all $j = 1, ..., n$, $\mathbb{K}_{j,n}$ is the number of consecutive flat trades before instant j, namely

$$\mathbb{K}_{j,n} = \min\left\{k \in \{0,...,j\} \mid \mathbb{B}_{j,n} = 1, \mathbb{B}_{j-1,n} = 1, ..., \mathbb{B}_{j-k+1,n} = 1, \mathbb{B}_{j-k,n} = 0\right\}.$$

Now, note that

$$\mathbb{K}_{j,n}=\begin{cases}0 & \text{with probability } 1-p_n\\ 1 & \text{with probability } \left(1-p_n\right)p_n\\ 2 & \text{with probability } \left(1-p_n\right)\left(p_n\right)^2\\ \vdots & \vdots\\ j-1 & \text{with probability } \left(1-p_n\right)\left(p_n\right)^{j-1}\\ j & \text{with probability } \left(p_n\right)^j .\end{cases}$$

Hence, we have

$$\begin{aligned}\mathbb{P}(\mathrm{R}_{j,j+1}|\mathrm{ES}^c_{j,j+1}) &= \sum_{k=0}^{j}\mathbb{P}\left[\mathbb{K}_{j,n}=k\right]\frac{1}{d}\int_{-d/2}^{d/2}\mathbb{P}\left(\overline{P}^-_{j\Delta_n}<P_{(j+1)\Delta_n}<\overline{P}^+_{j\Delta_n}|P_{(j-k)\Delta_n}=\overline{P}_{j\Delta_n}+x\right)dx\\ &= \sum_{k=0}^{j-1}\left(1-p_n\right)p_n^k\frac{1}{d}\int_{-d/2}^{d/2}\mathbb{P}\left(\overline{P}^-_{j\Delta_n}<P_{(j+1)\Delta_n}<\overline{P}^+_{j\Delta_n}|P_{(j-k)\Delta_n}=\overline{P}_{j\Delta_n}+x\right)dx\\ &\quad+p_n^j\frac{1}{d}\int_{-d/2}^{d/2}\mathbb{P}\left(\overline{P}^-_{j\Delta_n}<P_{(j+1)\Delta_n}<\overline{P}^+_{j\Delta_n}|P_0=\overline{P}_{j\Delta_n}+x\right)dx.\end{aligned}$$

In sum

$$\begin{aligned}\mathbb{P}(\mathrm{TS}_{j,j+1}) &= p_n+\sum_{k=0}^{j-1}\left(1-p_n\right)^2p_n^k\frac{1}{d}\int_{-d/2}^{d/2}\mathbb{P}\left(\overline{P}^-_{j\Delta_n}<P_{(j+1)\Delta_n}<\overline{P}^+_{j\Delta_n}|P_{(j-k)\Delta_n}=\overline{P}_{j\Delta_n}+x\right)dx\\ &\quad+\left(1-p_n\right)p_n^j\frac{1}{d}\int_{-d/2}^{d/2}\mathbb{P}\left(\overline{P}^-_{j\Delta_n}<P_{(j+1)\Delta_n}<\overline{P}^+_{j\Delta_n}|P_0=\overline{P}_{j\Delta_n}+x\right)dx.\end{aligned}$$

The conditional probabilities on the right-hand side are known in closed-form for all j (see Eq. (2)). The statement is conditional (on time j information) but it also holds unconditionally. Hence, a natural estimator $\widehat{p}_n$ of p_n is one which minimizes the distance between

$$Z_{0,n}:=\frac{1}{nT_k}\sum_{j=0}^{nT_k-1}\mathbf{1}_{\{\ln\overline{P}_{(j+1)\Delta_n}-\ln\overline{P}_{j\Delta_n}=0\}}$$

and

$$
\begin{aligned}
Z(p_n) \; := \; & p_n + \frac{1}{nT_k}\sum_{j=1}^{nT_k}\sum_{k=0}^{j-1}\left[\left(\operatorname{erf}\left(\frac{d}{\sqrt{2\widehat{\sigma}^2_{j\Delta_n}\overline{P}^2_{j\Delta_n}(k+1)\Delta_n}}\right)+\right.\right. \\
& \left.\left.\sqrt{\frac{2}{\pi}}\left(e^{-\frac{d^2}{2\widehat{\sigma}^2_{j\Delta_n}\overline{P}^2_{j\Delta_n}(k+1)\Delta_n}}-1\right)\frac{\widehat{\sigma}_{j\Delta_n}\overline{P}_{j\Delta_n}\sqrt{(k+1)\Delta_n}}{d}\right)(1-p_n)^2\, p_n^k\right] \\
& +\frac{1}{nT_k+1}\sum_{j=1}^{nT_k}\left(\operatorname{erf}\left(\frac{d}{\sqrt{2\widehat{\sigma}^2_{j\Delta_n}\overline{P}^2_{j\Delta_n}(j+1)\Delta_n}}\right)+\right. \\
& \left.\sqrt{\frac{2}{\pi}}\left(e^{-\frac{d^2}{2\widehat{\sigma}^2_{j\Delta_n}\overline{P}^2_{j\Delta_n}(j+1)\Delta_n}}-1\right)\frac{\widehat{\sigma}_{j\Delta_n}\overline{P}_{j\Delta_n}\sqrt{(j+1)\Delta_n}}{d}\right)(1-p_n)\, p_n^j,
\end{aligned}
$$

i.e.,

$$
\widehat{p}_n = \min_{p_n}\left|Z(p_n) - Z_{0,n}\right|. \tag{5}
$$

In Figure 6, we plot $\widehat{p}_n$, averaged across stocks, as a function of frequency. Consistent with the model (c.f., Eq. (2)), the probability of excess staleness is frequency-specific and, as expected, increasing with the sampling frequency. According to the model, there could be clustering in excess staleness, but the assumed (in Eq. (3)) dependence structure of the trading indicators (i.e., the Bernoulli variates) is such that their empirical averages converge to the probability of excess staleness at the highest frequency of observation (i.e., p_∞), when the observation frequency increases without bound (as $n \to \infty$). If we were to assume that the highest frequency is 30 seconds, then Figure 6 would imply that $p_\infty > 20\%$. Equivalently, at 30 seconds, one in five return observations is stale, a sizable magnitude. The corresponding probability of idleness (c.f., zeros associated with zero volumes) is marginally lower at 30 seconds, around 5% at 1 minutes, and lower than 1% at 2 minutes. This probability decays at a quicker pace than the probability of excess staleness. The increasing wedge between them suggests that, at low frequencies, all observed zeros occur in the presence of volumes which are strictly non-zero, albeit small. This observation is consistent with, e.g., Figure 3, Panel C.

We now return to the relation between staleness and volumes, something on which we began reporting in Figure 3 from a different perspective. Specifically, we evaluate the cross-sectional relation between the average probability of excess staleness, $\widehat{p}_n$, and dollar volumes. We do so in Figure 5, Panel A (30-second frequency), Panel B (1-minute frequency) and Panel C (2-minute frequency). At all frequencies, the correlations are negative, implying a lower likelihood of staleness in conjunction with higher volumes, and remarkably large (-90%, -80.3%, and -64.7%, respectively).

Should price discreteness induce a large percentage of spurious zeros, i.e., zeros associated with relatively large volumes, we would expect total staleness to display substantially lower (in absolute value) correlation with respect to volumes than excess staleness. The corresponding results are

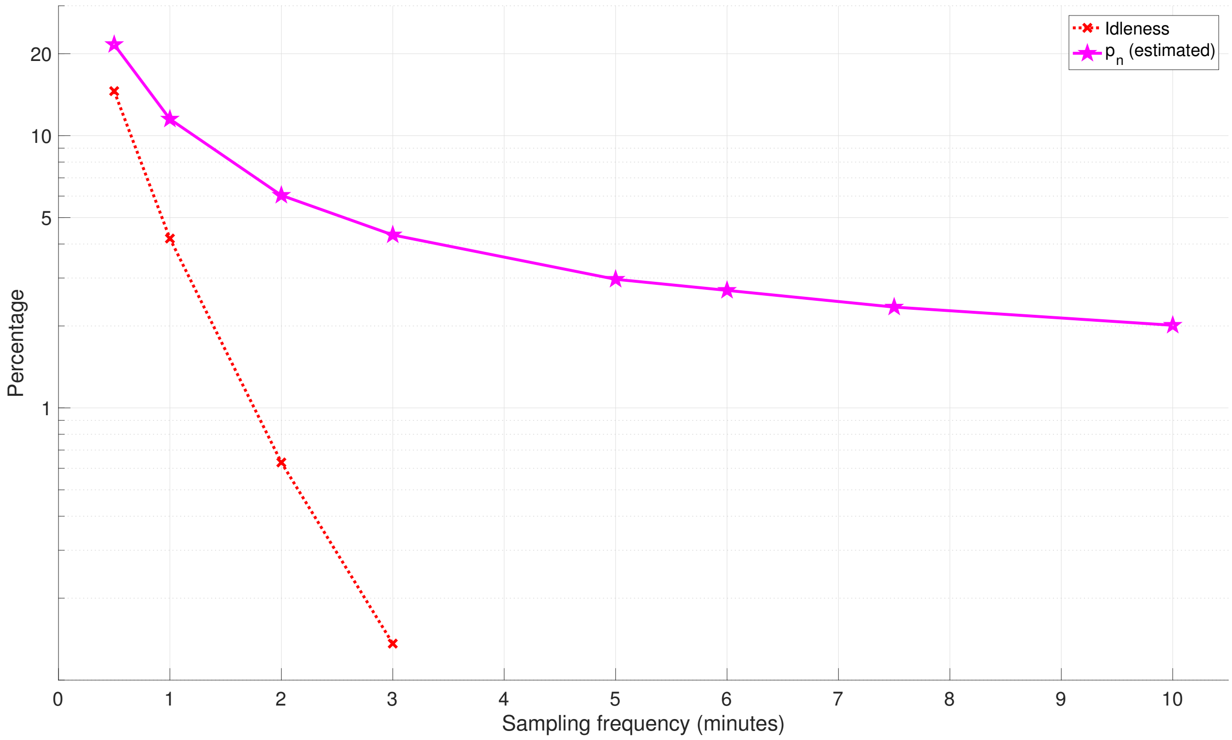

Figure 6: We report the average probability of excess staleness (purple line) and the average probability of idleness (red line) as a function of the sampling frequency. We consider sampling frequencies equal to 30 seconds and 1 minute through 10 minutes. The probability of excess staleness is defined as in Eq. (5). The probability of idleness is the frequency of zero returns without volume.

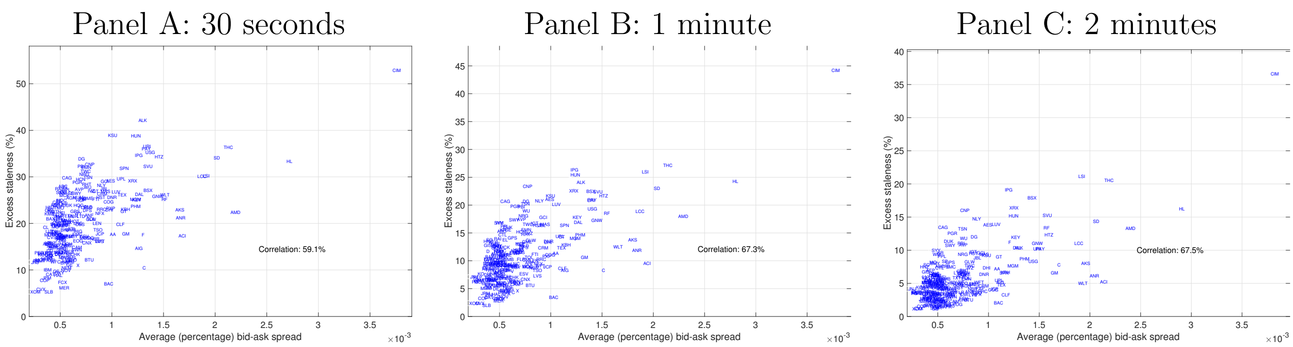

Figure 7: We plot the average probability of excess staleness versus the average logarithmic bid/ask spread over three frequencies: 30 seconds (Panel A), 1 minute (Panel B) and 2 minutes (Panel C).

reported in Panel D (30 seconds), Panel E (1 minute) and Panel F (2 minutes) of Figure 6. While, in agreement with the confounding effect of price discreteness, we do witness a reduction in correlation (particularly at high frequencies), the relation between staleness (total, this time) and volumes continues to be strong (-70% at 30 seconds, -60% at 1 minute, and -51.6% at 2 minutes, respectively). Once more, rounded up prices which do not move because of small volumes are not in contradiction with our documented negative relation between staleness and volumes. In this sense, a (possibly large) component of price discreteness is, just like in the case of excess staleness, a reflection of slow trading.

7 The economics of staleness: further discussion

As stressed above, the reported findings may be suggestive of a positive relation between illiquidity and excess staleness with volumes representing the mediating variable. In order to verify this conjecture, we employ the most classical liquidity measure, namely traded logarithmic bid/ask spreads. We do so keeping in mind the well-known limitations of the quoted bid/ask spread as a comprehensive measure of the trade cost faced by market participants.

For the firms in our sample, Figure 7 reports the cross-sectional relation between the average probability of excess staleness and average logarithmic bid/ask spreads associated with three frequencies, 30 seconds (Panel A), 1 minute (Panel B) and 2 minutes (Panel C). In a nutshell, firms with more excess staleness have, on average, larger spreads. The magnitudes of the (positive) correlations are noteworthy (59% at 30 seconds, 67.3% at 1 minute, and 67.5% at 2 minutes).

Let us now focus on a typical firm with, say, more excess staleness. While the company has larger bid/ask spreads on average, it is an empirical question as to whether the level of staleness adjusts to the size of the spreads. Said differently, the firm may simply have more zeros than a firm with

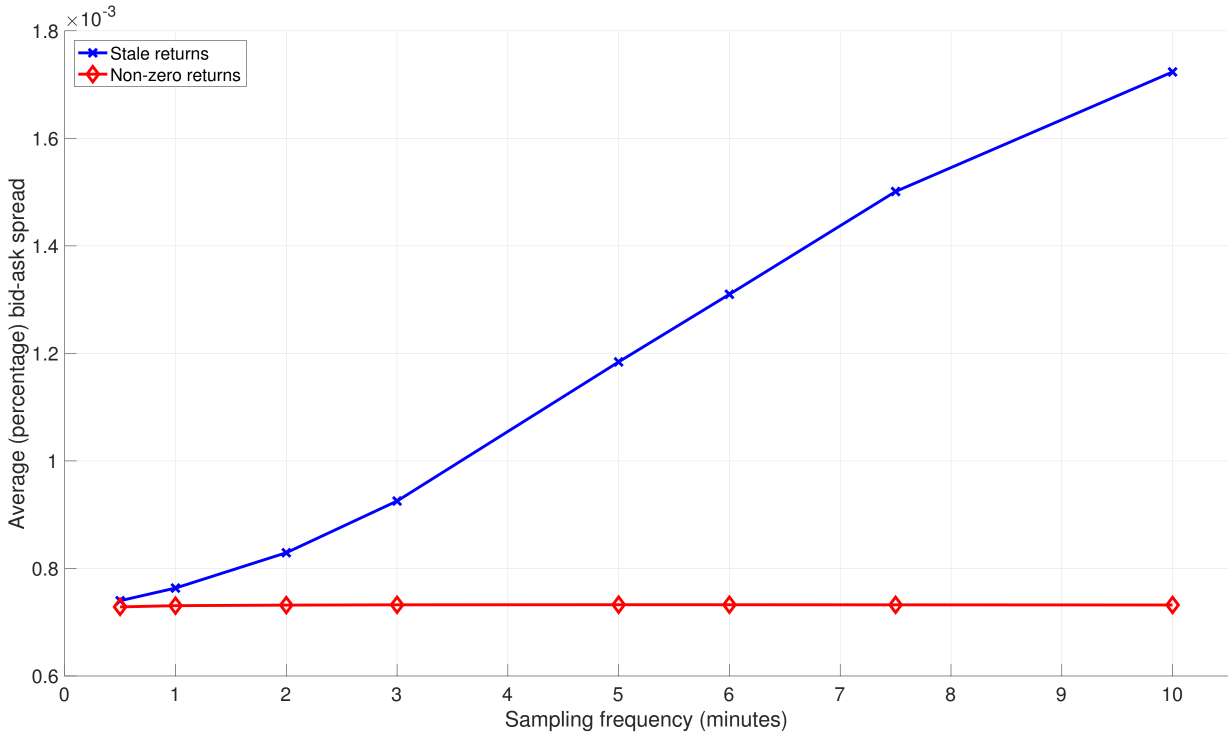

Figure 8: We plot average logarithmic bid/ask spreads associated with zero returns (blue line) and non-zero returns (red line) across sampling frequencies. The frequencies are 30 seconds and 1 minute through 10 minutes.

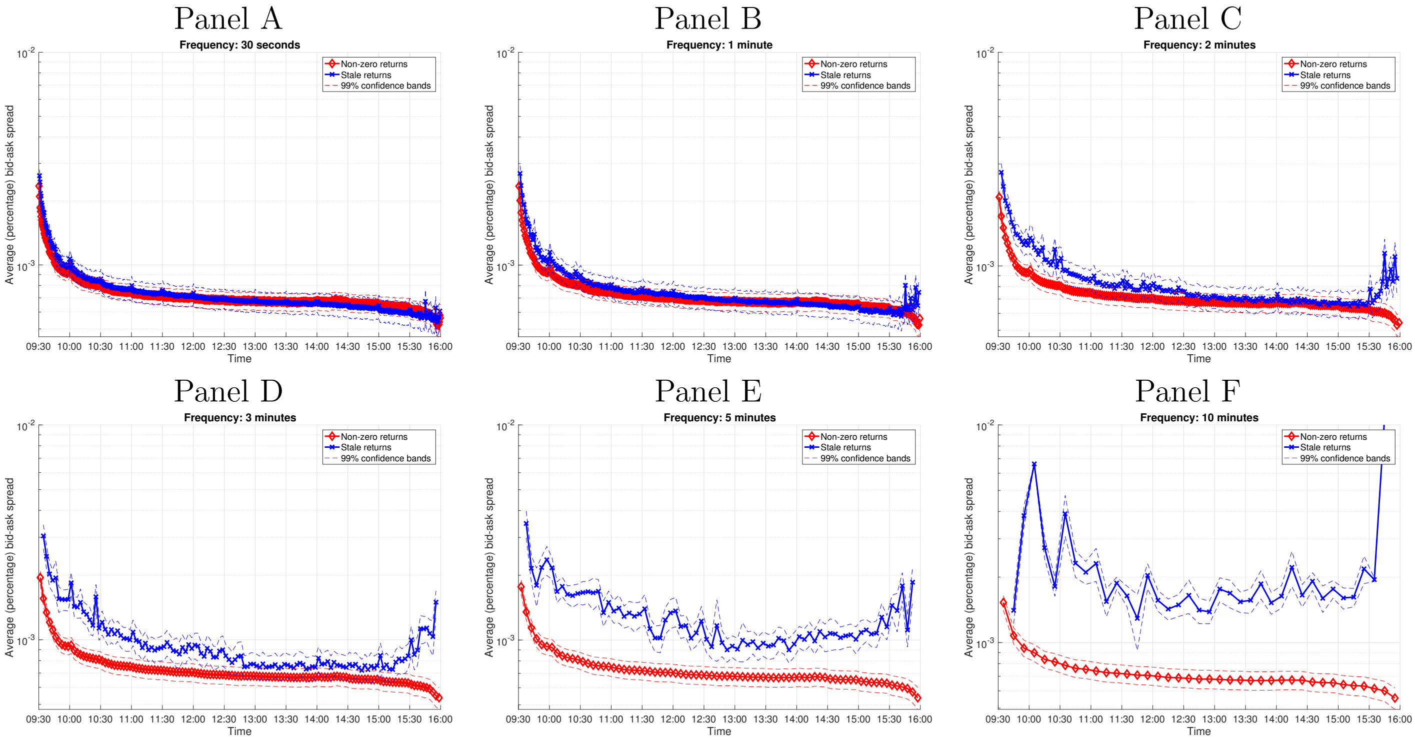

Figure 9: We plot average logarithmic bid/ask spreads associated with zero returns (blue lines) and non-zero returns (red lines) across times of the day and frequencies. The frequencies are 30 seconds, 1 minutes, 2 minutes, 3 minutes, 5 minutes, and 10 minutes.

lower spreads, but equally large spreads associated with zeros and non zeros.

Figure 8 reports the average size of the logarithmic bid/ask spreads associated with zeros (blue line) and non-zero returns (red line) for various frequencies (30 seconds and 1 minute through 10 minutes). Figure 9 visualizes the intra-daily structure in the average bid/ask spreads associated with zeros (blue line) and non-zero returns (red line) over 6 frequencies (30 seconds, 1 minute, 2 minutes, 3 minutes, 5 minutes and 10 minutes). In both cases, we relate the mean of the logarithmic bid-ask spreads over the specific time horizon to zero and non-zero returns over the same horizon.

The relation between the spreads and zeros is frequency-specific. We report a monotonic increase (over frequencies) in the average size of the spreads associated with zeros. The size of the spreads associated with non-zero returns is, instead, stable. Over the 30-second horizon and similar high frequencies, the spreads are equally large for zeros and non zeros (on average, Figure 8, and intra-daily, Figure 9). The wedge between spreads associated with zeros and spreads associated with non-zero returns increases with reductions in sampling frequency (Figure 8). So, does the intra-daily wedge (Figure 9).

This empirical evidence supports the economic logic discussed in Section 4. Other things equal, it is consistent with the baseline model in Bandi et al. (2017): informed traders, assumed to trade on (or near) the spread when the absolute difference between quoted midpoints of bid/ask spreads and fundamental values is larger than the half spread plus, e.g., funding costs, will trade less in the presence of larger spreads. Said differently, when faced with a larger "no-trade region," as determined by larger execution costs, informed traders are more likely to opt out of trading. It is also consistent with the third extension of the baseline model in Bandi et al. (2017): uninformed (or noise) traders may be reluctant to trade if the *absolute* size of the spreads is excessively large.

In sum, long spells of higher (on average) execution costs, as represented here by higher quoted bid/ask spreads, have an impact on traders' willingness to transact. Conversely, it would appear that, in order to have zeros at lower frequencies, the quoted spreads should be relatively large. The longer the horizon of a zero return, the larger the spreads need to be to induce prolonged traders' inactivity.

8 Conclusions

This paper provides empirical foundations for the relevance of staleness in high-frequency asset prices. Institutional effects, such as price discreteness, do not represent a first-order determinant of staleness, a stylized fact shown to be driven by volume and the magnitude of execution costs.

Our evidence supports an alternative data generating processes for asset prices, one which deviates in important ways from classical semi-martingale modeling with (locally) Brownian shocks and, instead, allows for clusters of inactivity.

We expect the presence, number and dynamic properties of zeros to be informative about the features of the price formation mechanism, from the extent of asymmetric information, to the cost of trading, to the dynamic evolution of unobserved fundamental values. The structural analysis in Bandi et al. (2017), justified by the empirical work in this paper, is only a first step in this area.

9 Acknowledgments

We thank conference participants at the Third International Workshop in Financial Econometrics, Arraial d'Ajuda Ecoresort, Brazil, October 8-10, for discussion. We are particularly indebted to the discussant, Jantje Soenksen, for her comments and suggestions. Her discussion and communications with us have lead to a sharper treatment.

References

Bandi, F., D. Pirino, and R. Renò (2017). EXcess Idle Time. *Econometrica*. Forthcoming.

Broussard, J. P. and A. L. Nikiforov (2013). Human bias in algorithmic trading. Working paper.

Delattre, S. and J. Jacod (1997). A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors. *Bernoulli*, 1–28.

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

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

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica 53*(6), 1315–1335.

Tauchen, G. and M. Pitts (1983). The price variability-volume relationship on speculative markets. *Econometrica 51*, 485–505.