Portfolio cross-autocorrelation puzzles*

Dan Bernhardt†
Department of Economics
University of Illinois
1206 South Sixth Street
Champaign IL 61820

Ryan J. Davies‡
Finance Division
Babson College
Tomasso Hall
Babson Park MA 02457

Preliminary: November 1, 2005

Abstract

This paper investigates the driving forces underlying lead-lag cross-autocorrelations in daily portfolio returns. By contrasting autocorrelation patterns using portfolio returns based on trade prices before an arbitrary point in the trading day with those using returns based on prices after, we isolate the impact of nonsynchronous trading and conclude decisively that its impact on portfolio correlation patterns is negligible. We also reject the possibility that time-varying returns, contemporaneous correlations or discreteness in pricing underlies the portfolio autocorrelation patterns. Rather, the intradaily evolution of portfolio autocorrelations reflect inefficient pricing of infrequently-traded stocks—prices of less active stocks appear not to incorporate some of the recent information that is already contained in the prices of more active stocks. Portfolio autocorrelations rise systematically when we calculate returns using later times in the day, indicating that prices of infrequently-traded stocks grow increasingly stale over the trading day.

JEL Classification: G12, G14.

Keywords: portfolio returns, autocorrelation, non-synchronous trading, market efficiency.

---

*We are grateful for comments received from seminar participants at Babson, Illinois, Reading, Warwick, HEC Montréal, and the Northern Finance Association meetings. Bernhardt acknowledges financial support from NSF grant SES-0317700. Davies acknowledges support from the Babson College Board of Research.

†Tel: +1-217-244-5708. email: danber@ragnar.econ.uiuc.edu

‡Tel: +1-781-239-5345. Fax: +1-781-239-5004. email: rdavies@babson.edu

Portfolio cross-autocorrelation puzzles

This paper investigates the driving forces underlying lead-lag cross-autocorrelations in daily portfolio returns. By contrasting autocorrelation patterns using portfolio returns based on trade prices before an arbitrary point in the trading day with those using returns based on prices after, we isolate the impact of nonsynchronous trading and conclude decisively that its impact on portfolio correlation patterns is negligible. Instead, the patterns reflect inefficient pricing of infrequently-traded stocks—prices of less active stocks appear not to incorporate some of the recent information that is already contained in the prices of more active stocks. Portfolio autocorrelations rise systematically when we calculate returns using later times in the day, indicating that prices of infrequently-traded stocks grow increasingly stale over the trading day.

JEL Classification: G12, G14.

Keywords: portfolio returns, autocorrelation, non-synchronous trading, market efficiency.

1 Introduction

It is well-established that short-horizon portfolio returns are both significantly autocorrelated and highly cross-serially correlated (see, for example, Boudoukh, et al. (1994)). Proposed explanations of these empirical regularities include: market inefficiency, market maker inventory control, transaction costs, short sale constraints, price discreteness, time-varying expected returns, and nonsynchronous trading. Still, the "puzzle" remains largely unresolved—there is much disagreement about the relative importance of each of these explanations.

Unlike previous studies, we are able to isolate the impact of nonsynchronous trading on portfolio autocorrelation patterns and to uncover the extent to which delayed incorporation of information into the prices of some stocks drives portfolio autocorrelations. By contrasting autocorrelations for returns calculated using the last transaction before a given point in the trading day versus those calculated using the first transaction after, we provide sharp theoretical bounds on the autocorrelations that can be due to nonsynchronous trading. This allows us to conclude decisively that the impact of nonsynchronous trading is negligible.

We then derive theoretically how variations in information arrival rates, delayed incorporation of information into prices and trading frequencies affect autocorrelation patterns. Numerically, we uncover strong evidence that autocorrelation patterns reflect inefficient pricing of infrequently-traded stocks—prices of less active stocks appear not to incorporate recent information contained in the prices of more active stocks. In particular, portfolio autocorrelations rise several fold when we calculate returns using later times in the day, indicating that prices of infrequently-traded stocks grow increasingly stale over the trading day.

Our findings have important implications for a wide variety of market participants. For instance, our results suggests that mutual funds which value purchases and redemptions by calculating net asset values (NAV) using end-of-day prices may, in fact, be using inaccurate prices, particularly for smaller, less frequently traded stocks.[1] The same also applies for derivatives which are settled based on closing prices. As well, trading strategies based on

[1]See Chalmers, Edelen, Kadlec (2001) for a description of this problem.

trading at the close will need to balance the benefits of greater liquidity against these stale prices. Finally, to the extent which "staleness" is driven by limit orders which have not been updated, our findings have implications for market design.[2]

Further analysis precludes other possible explanations of portfolio autocorrelations. Portfolio autocorrelation patterns are the same whether we use prices or quotes, indicating that discreteness in pricing does not drive the autocorrelations. Further, contemporaneous portfolio correlations do not vary over the trading day, while portfolio autocorrelations do, indicating that the phenomena are distinct, and do not reflect time-varying expected returns. In sum, our numerical findings strongly indicate that the intra-day evolution of pricing inefficiencies—price staleness—drives the secular rise over the trading day in portfolio autocorrelations.

To begin, let us see how we identify the impact of nonsynchronous trading. The potential qualitative impact of nonsynchronous trading on portfolio autocorrelations can be gleaned by considering the standard practice of using closing prices to calculate returns. Consider two stocks, X and Y. Suppose that stock X consistently trades later in the day than stock Y. If significant economic news arrives after stock Y has finished trading, but before stock X's last trade, then the closing price of stock X will reflect this information, but the closing price of stock Y will not. When the stocks resume trading the next day, stock Y's price will be updated to reflect this information. The consequence is that stock X's returns will lead (or predict) returns of stock Y. This cross-serial correlation across stocks contributes to the observed portfolio return autocorrelation.

Our key insights are to recognize that (i) we can compute daily returns using arbitrary points in the day; and (ii) the impact of nonsynchronous trading on autocorrelation relationships is *reversed* if we compute daily returns using the first trade *after* the arbitrary time, rather than the last trade *before* that moment. This permits us to isolate the impact of nonsynchronous trading on portfolio autocorrelations.

[2]Cohen et al. (1986) describe three types of price-adjustment delay: (i) transaction price adjustment lag quotation price adjustments; (ii) specialists-dealers impede quotation price adjustments; and (iii) individual wait-to-trade delays.

To see this, suppose that we compute daily (24 hour) returns using the last trade *before* noon of each trading day. Then trades of frequently-traded stocks will tend to occur closer to noon than will trades of infrequently-traded stocks. As a result, lagged stock prices of frequently-traded stocks will tend to contain economic news that arrived after the last trade (prior to noon) of infrequently-traded stocks. As a result the impact of nonsychronous trading when we compute returns using the last trade before noon should be similar to that when we compute returns using closing prices.

But now suppose that we use the first trade *after* noon to compute returns. Then the impact of nonsynchronous trading is reversed precisely because the first trade after noon of infrequently-traded stocks tends to occur after that of the first trade of frequently-traded stocks: if stock prices reflect all market information, then prices of infrequently-traded stocks will tend to contain economic news that arrived after the frequently-traded stocks traded.

Empirically, we find that *even* when 24-hour returns are computed using the first trade *after* a moment in time, lagged returns of large-cap/frequently-traded stocks better predict current returns on small-cap/infrequently-traded stocks than the converse ($\mathrm{corr}(r^L_{t-1}, r^S_t)$ $>>$ $\mathrm{corr}(r^S_{t-1}, r^L_t)$). This means that pricing of infrequently-traded stocks *must* reflect some inefficiency—when prices of infrequently-traded stocks are determined, they must not incorporate all of the recent information contained in the prices of frequently-traded stocks. This is because the impact of nonsynchronous trading in the absence of delayed incorporation of information gives rise to the opposite correlation pattern.

To isolate the slight impact of nonsynchronous trading, we must control for differences in information content across portfolios. We do this by calculating the before-after *difference* in portfolio autocorrelations for a *fixed* pair of portfolios. We consistently find evidence that the impact of nonsynchronous trading exists, but that it is very small in magnitude. Specifically, the correlation of lagged large-cap portfolio returns with current small-cap portfolio returns is slightly higher using the last trade before than using the first trade after; while the correlation of lagged small cap portfolio returns with current large-cap portfolio returns is slightly higher using the first trade after, than using last trade before. However, these differences, which are roughly twice the impact of nonsynchronous trading, are on the order of 0.005. This result is robust: it holds no matter (i) which point in time we use to compute before-after return differences, (ii) whether we compute returns using trades or quotes, and (iii) whether we construct portfolios by market capitalization or trading frequency.

We next investigate the consequences of delayed incorporation of extant information into the prices of less active stocks. No matter whether we compute returns using the last trade before a moment, or the first trade after, the small-capitalization portfolio own-autocorrelation is always large and positive, exceeding the cross-portfolio lead-lag autocorrelations, and greatly exceeding large-capitalization portfolio own-autocorrelation. This can be reconciled if information is impounded more slowly into some small-capitalization stocks than into others; but information is impounded uniformly more quickly into large capitalization stocks.

We then provide strong evidence that some stock prices grow increasingly stale over the trading day. Specifically, both own- and cross-autocorrelations rise dramatically throughout the trading day, before decreasing slightly at the close. For example, the own-autocorrelation for small-cap portfolio returns rises from about 0.1 at 1100h to about 0.2 at 1430h and to 0.218 at close, a more than two-fold increase. The cross-autocorrelation between current large-cap portfolio returns and lagged small-cap portfolio returns rises even more dramatically from about 0.017 at 1100h to about 0.120 at 1430h before declining to 0.093 at close. These intra-daily differences in cross-autocorrelations are all highly significant.

Crucially, because we control for the impact of trading frequencies, this strong secular rise in autocorrelations must be due to delayed incorporation of information. Our evidence suggests that during the overnight period, investors gather information and stock prices "catch up" but as the trading day progresses, the incorporation of new information into some stock prices becomes increasingly lagged. The different patterns observed at the close are consistent with large stock prices getting "fresher" at the close. In turn, this magnifies the predictive power of lagged large-cap portfolio returns on current small-cap portfolio returns and is not representative of the autocorrelation patterns observed at other times in the day.

The remainder of the paper is organized as follows. Section 2 reviews related research.

Section 3 provides a simple theoretical model of the impact of nonsynchronous trading and delayed incorporation of information for portfolio autocorrelations. Section 4 outlines the data. Section 5 provides the results. Section 6 concludes.

2 Literature Review

Conrad and Kaul (1988) claim that cross-autocorrelations are the result of time-varying expected returns. A variant of this explanation suggests that cross-autocorrelations are simply a restatement of portfolio autocorrelations and contemporaneous correlations (Hameed (1997)). According to this explanation, once account is taken of portfolio autocorrelations, portfolio cross-autocorrelations should disappear. The second group of explanations (Boudoukh et al. (1994)) suggest that portfolio autocorrelations and cross-autocorrelations are due to market microstructure biases such as thin trading or discreteness in prices. A third explanation is that the lead-lag cross-autocorrelations are caused by the tendency of some stocks to adjust more slowly (under-react) to economy-wide information than others (Lo and MacKinlay (1990) and Brennan et al. (1993)).

Our paper provides insights into each of these possible explanations. Specifically, we show that the impact of thin trading (nonsynchronous trading) is minor, and the fact that autocorrelations patterns do not depend on whether we use trades or quote midpoints indicates that discreteness is unimportant. We find that contemporaneous correlation patterns do not vary over the trading day, but lead-lag correlations rise secularly over the trading day: Hameed's hypothesis cannot account for this intraday pattern. Rather, it appears that this intraday evidence can only be reconciled by inefficient pricing of some stocks.

Atchison et al. (1987) and Lo and MacKinlay (1990) derive explicit relations concerning the magnitude of autocorrelations caused by nonsynchronous trading. Both studies conclude that nonsynchronous trading can explain only a small portion of the autocorrelation that is observed. However, Boudoukh et al. (1994) argue that prior studies seriously understate the potential effects of nonsynchronous trading. For instance, Boudoukh et al.

show that portfolios with weekly autocorrelations of 0.07 under standard assumptions can have autocorrelations as high as 0.20 when the nonsynchronous trading framework allows for heterogeneity in both nontrading probabilities and security betas.

Kadlec and Patterson (1999) simulate the effects of nonsynchronous trading by sampling stock returns from a return-generating process using transactions data to obtain the precise time of each stock's last trade. Their simulated weekly portfolio returns exhibit autocorrelations that are about 25% of observed autocorrelations. However, our before-after decomposition reveals that nonsynchronous trading has only a small fraction of this impact.

Chordia and Swaminathan (2000) find that daily returns on high trading volume portfolios lead returns on low volume portfolios, controlling for firm size. While our theoretical decomposition shows that this could be reconciled by both nonsynchronous trading due to trading frequency differences, or by delayed incorporation of information into prices of low volume stocks; our empirical work indicates that it reflects the delayed incorporation of information.

Chan (1993) provides a model in which market makers observe noisy signals about their stocks and correct pricing errors by observing the past price changes in other stocks. This can generate delayed incorporation of information into stock prices, which is consistent with our empirical findings.

Researchers have studied many other aspects of portfolio autocorrelation patterns. Papers that explore autocorrelation patterns over longer intervals and longer lags, and the profitability of contrarian or momentum strategies include Jegadeesh and Titman (1995) and Lewellen (2002). Sias and Starks (1997), and Badrinath, et al. (1995) look at how trading by institutional traders in specific groups of stocks can affect portfolio autocorrelations. Other important papers on portfolio autocorrelation patterns include McQueen, Pinegar, and Thorley (1996), Mech (1993), and Bessembinder and Hertzel (1993).

The closest related work to ours is McInish and Wood (1991). Using a 1984 data tape of 1,400 NYSE stocks, they calculate intraday-to-intraday 24-hour returns at successive fifteen minute intervals throughout the trading day. They consider six equally-weighted indexes:

one comprised of all stocks and five comprised of stocks ranked by quintiles of average number of trades per day over the year. They show that the plot of first-order autocorrelation of daily index returns against time of day has a crude U-shaped pattern. For each stock, they calculate the average time from the last trade (which might occur prior to the beginning of the interval) to the interval end. Using these times, they argue that the observed intraday patterns in first-order autocorrelation of return indexes cannot be explained by trading delays.

Our analysis extends that of McInish and Wood (1991) in four key respects: (i) we compare correlation patterns before and after a particular point in time; (ii) we examine cross-correlation patterns, whereas McInish and Wood focus exclusively on own-autocorrelation patterns; (iii) we use a longer, more recent time-series data which allows for us to study the impact of market design changes; (iv) we use an advanced block of blocks bootstrap technique to establish the statistical significance of our findings.

3 Theory

We next develop how the impact of nonsynchronous trading on portfolio autocorrelations depends on the 24-hour window used to compute returns. We first consider a base case of identical stocks that differ only in their trading frequency and derive the lead-lag autocorrelation patterns for returns computed over different windows. We then allow for random trade arrival. Finally, we consider how outcomes are affected when prices of some stocks do not incorporate the latest valuation information.

The base framework: Consider two stocks, X and Y, that each have an identical end-of-date t value given by

$$V_t = V_{t-1} + \delta_t^0 + \delta_t^1 + \ldots + \delta_t^\ell + \delta_t^{\ell+1} + \ldots + \delta_t^c,$$

where, for simplicity, we assume that the innovations δ_t^j are independently distributed. Interpret δ_t^0 as an overnight innovation that arrives after the date $t-1$ market closes, δ_t^1 as the first innovation after the market opens and δ_t^c as the last innovation before market close.

Stocks X and Y are distinguished solely by the fact that X trades more frequently than Y. Specifically, we suppose that X trades before and after each innovation; while Y last trades just before δ_t^c has been realized, and Y first trades just after δ_t^1 has been realized. Then, if the innovations are immediately incorporated into prices, the closing price of X will be $p_t^{Xc} = V_t$, but the closing price of Y will not incorporate the last innovation so that $p_t^{Yc} = V_t - \delta_t^c$. In sharp contrast, the opening price of X will be $p_t^{Xo} = V_{t-1} + \delta_t^0$, but Y's opening price will incorporate the first innovation of the trading day, $p_t^{Yo} = V_{t-1} + \delta_t^0 + \delta_t^1$. See figure 1.

We now consider the implications of trading frequencies for the covariance in daily lead-lag price changes, i.e., $\mathrm{cov}(p_{t+1} - p_t, p_t - p_{t-1})$. Our exposition considers covariance patterns in daily price changes rather than returns because the qualitative insights are identical, and they are most easily presented via price changes.[3] When innovations are independently distributed, then no matter whether we use opening or closing prices, lagged change in own price is always uncorrelated with current changes in own price. The closing price of stock X reflects all of the day's innovations, so that $p_{t+1}^{Xc} - p_t^{Xc} = \delta_t^0 + \delta_t^1 + \ldots + \delta_t^c$. Thus,

$$\mathrm{cov}\left(p_{t+1}^{Xc} - p_t^{Xc}, p_t^{Xc} - p_{t-1}^{Xc}\right) = \mathrm{cov}\left(\delta_t^0 + \delta_t^1 + \ldots + \delta_t^c, \delta_{t-1}^0 + \delta_{t-1}^1 + \ldots + \delta_{t-1}^c\right) = 0.$$

Similarly, $p_{t+1}^{Yc} - p_t^{Yc} = \delta_{t-1}^c + \delta_t^0 + \delta_t^1 + \ldots + \delta_t^{c-1}$, so that

$$\mathrm{cov}\left(p_{t+1}^{Yc} - p_t^{Yc}, p_t^{Yc} - p_{t-1}^{Yc}\right) = \mathrm{cov}\left(\delta_{t-1}^c + \delta_t^0 + \ldots + \delta_t^{c-1}, \delta_{t-2}^c + \delta_{t-1}^0 + \ldots + \delta_{t-1}^{c-1}\right) = 0.$$

Matters are very different when we compute the cross-stock lead-lag autocorrelations. Using closing prices, the lagged price change of stock X contains the innovation δ_t^c that enters the current price change of stock Y. As a result,

$$\begin{aligned}\mathrm{cov}\left(p_{t+1}^{Yc} - p_t^{Yc}, p_t^{Xc} - p_{t-1}^{Xc}\right) &= \mathrm{cov}\left(\delta_{t-1}^c + \delta_t^0 + \delta_t^1 + \ldots + \delta_t^{c-1}, \delta_{t-1}^0 + \delta_{t-1}^1 + \ldots + \delta_{t-1}^c\right) \\ &= \mathrm{var}(\delta_{t-1}^c) > 0.\end{aligned}$$

This is the well-understood impact of nonsynchronous trading: changes in the closing prices of frequently-traded stocks contain information that has yet to be incorporated in the closing prices of infrequently-traded stocks, leading to a positive cross-autocorrelation.

[3]Alternatively, this presentation captures returns if innovations are multiplicative and prices are in logs.

Conversely, lagged price changes in the infrequently-traded stock, Y are uncorrelated with current price changes in X, because there is no information overlap,

$$\text{cov}\left(p_{t+1}^{Xc} - p_t^{Xc}, p_t^{Yc} - p_{t-1}^{Yc}\right) = \text{cov}\left(\delta_t^0 + \delta_t^1 + \ldots + \delta_t^c, \delta_{t-2}^c + \delta_{t-1}^0 + \delta_{t-1}^1 + \ldots + \delta_{t-1}^{c-1}\right) = 0.$$

We now show that the impact of nonsynchronous trading is reversed if we use opening prices to compute changes. The cross-autocorrelation patterns are reversed *precisely* because infrequently-traded stock Y's first trade occurs after that for X. As a result, lagged prices changes in Y contain information about δ_t^1 that will enter the current price change in X. Hence, lagged price changes in Y are positively correlated with current price changes in X, but not conversely,

$$\text{cov}\left(p_{t+1}^{Xo} - p_t^{Xo}, p_t^{Yo} - p_{t-1}^{Yo}\right) = \text{var}\left(\delta_t^1\right); \text{ and } \text{cov}\left(p_{t+1}^{Yo} - p_t^{Yo}, p_t^{Xo} - p_{t-1}^{Xo}\right) = 0!$$

Indeed, if there is more information arrival at open than at close, i.e., if $\text{var}\left(\delta_t^1\right) > \text{var}\left(\delta_t^c\right)$, then this relationship should be even stronger than the standard lead-lag autocorrelations documented using closing prices.

Because information arrival rates may differ at open and close, i.e., $\text{var}(\delta_t^1) \neq \text{var}(\delta_t^c)$, inference about the impact of non-synchronous trading using only opening and closing prices may be difficult. However, we can control for information arrival rates if we instead compute returns using the last transaction before and the first transaction after a specified time ℓ in the trading day. Using the last transaction before time ℓ is akin to using closing prices—the last transaction of the frequently-traded stock before time ℓ will tend to be after the infrequently-traded stock, so lagged returns of the frequently-traded stock will contain information about current returns of the infrequently-traded stock. Conversely, using the first transaction after a moment in time, lagged returns in the infrequently-traded stock should predict current returns in the frequently-traded stock.

To make this explicit, suppose that δ_t^ℓ occurs before time ℓ, and is incorporated into the time ℓ price of stock X, but not of stock Y; and that the first transaction after time ℓ in stock X occurs before $\delta_t^{\ell+1}$ is realized, but that the first transaction in infrequently-traded stock Y occurs after (see figure 2). Then

$$\operatorname{cov}\left(p_{t+1}^{Yb\ell}-p_t^{Yb\ell},p_t^{Xb\ell}-p_{t-1}^{Xb\ell}\right)=\operatorname{var}\left(\delta_t^{\ell}\right); \quad \text{and} \quad \operatorname{cov}\left(p_{t+1}^{Yb\ell}-p_t^{Yb\ell},p_t^{Xb\ell}-p_{t-1}^{Xb\ell}\right)=0,$$

where the index $b\ell$ denotes the last transaction **b**efore time ℓ. Conversely if we used the first transaction price **a**fter ℓ, the lead-lag pattern is reversed:

$$\operatorname{cov}\left(p_{t+1}^{Ya\ell}-p_t^{Ya\ell},p_t^{Xa\ell}-p_{t-1}^{Xa\ell}\right)=0; \quad \text{and} \quad \operatorname{cov}\left(p_{t+1}^{Xa\ell}-p_t^{Xa\ell},p_t^{Ya\ell}-p_{t-1}^{Ya\ell}\right)=\operatorname{var}\left(\delta_t^{\ell+1}\right).$$

As long as information arrival just before ℓ is essentially the same as that just after ℓ, so that $\operatorname{var}(\delta_t^{\ell+1}) \sim \operatorname{var}(\delta_t^{\ell})$, the lead-lag relationships should be exactly reversed. Thus, the key force driving the impact of nonsynchronous trading on the lead-lag return correlation pattern is whether prices are computed using the last price before a moment in time (as at close), or the first price after a moment in time (as at open).

Finally, this example supposes that stock Y trades only slightly less often than stock X. Were Y to trade even less frequently, more innovations arrived between trades, raising the magnitude of the autocorrelations, but preserving otherwise the qualitative pattern.

Portfolios: In practice, the last trade of an infrequently-traded stock sometimes takes place after that of frequently-traded stocks. To determine how this affects autocorrelations, consider now portfolios of infrequently- and frequently-traded stocks, that are otherwise identical. Suppose that fraction ρ_f of **f**requently-traded stocks trade after innovation δ_t^{ℓ} is realized, but before time ℓ; and for the remaining fraction $1-\rho_f$, the last trade occurs before δ_t^{ℓ} is realized. In contrast, suppose that fraction $\rho_i < \rho_f$ of **i**nfrequently-traded stocks trade after innovation δ_t^{ℓ} is realized, but before time ℓ; and for the remaining fraction $1-\rho_i$, the last trade occurs after $\delta_t^{\ell-1}$, but before δ_t^{ℓ} is realized.

Thus, $\rho_f - \rho_i$ measures the degree to which frequently-traded stocks are more likely than infrequently-traded stocks to trade after an innovation. Because large capitalization stocks tend to trade more frequently than small capitalization stocks, one interpretation is that portfolios of frequently- and of infrequently-traded stocks are analogous to portfolios of large and of small capitalization stocks, respectively.

We denote the average price change for a very large portfolio of N **f**requently-traded stocks using the last price **b**efore time ℓ by

$$\Delta p_t^{fb\ell} = \frac{1}{N} \sum_{n=1}^{N} \left[p_t^{fb\ell}(n) - p_{t-1}^{fb\ell}(n) \right].$$

For this portfolio of **f**requently-traded stocks, the impact of non-synchronous trading for its own lead-lag covariance is

$$\text{cov}\left(\Delta p_t^{fb\ell}, \Delta p_{t-1}^{fb\ell}\right) = \rho_f(1-\rho_f)N\text{var}(\delta_t^{\ell}). \tag{1}$$

So, too, if fraction ρ_f of **f**requently-traded stocks also trade after time ℓ, but before innovation $\delta_t^{\ell+1}$ is realized, then using the first transaction **a**fter ℓ to compute the own lead-lag covariance for the portfolio frequently-traded stocks yields

$$\text{cov}\left(\Delta p_t^{fa\ell}, \Delta p_{t-1}^{fa\ell}\right) = \rho_f(1-\rho_f)N\text{var}(\delta_t^{\ell+1}).$$

The analogous lead-lag covariances using the portfolio of **i**nfrequently traded stocks are

$$\text{cov}\left(\Delta p_t^{ib\ell}, \Delta p_{t-1}^{ib\ell}\right) = \rho_i(1-\rho_i)N\text{var}(\delta_t^{\ell}) \quad \text{and} \quad \text{cov}\left(\Delta p_t^{ia\ell}, \Delta p_{t-1}^{ia\ell}\right) = \rho_i(1-\rho_i)N\text{var}(\delta_t^{\ell+1}).$$

If $\text{var}(\delta_t^{\ell+1}) \sim \text{var}(\delta_t^{\ell})$, own portfolio return lead-lag correlations should be essentially the same no matter whether we use the last trade before a particular time or first trade after. However, absent knowing ρ_f and ρ_i, we cannot say whether the lead-lag covariance should be stronger for frequently- or infrequently-traded stocks.

What about the cross-portfolio lead-lag covariances? If we compute returns using the last transaction **b**efore ℓ, then the portfolio of lagged price changes for **f**requently-traded stocks will co-vary more strongly with current price changes for **i**nfrequently-traded stocks, than its opposite counterpart. That is, because $\rho_f > \rho_i$,

$$\text{cov}\left(\Delta p_t^{ib\ell}, \Delta p_{t-1}^{fb\ell}\right) = \rho_f(1-\rho_i)N\text{var}(\delta_t^{\ell}) > \rho_i(1-\rho_f)N\text{var}(\delta_t^{\ell}) = \text{cov}\left(\Delta p_t^{fb\ell}, \Delta p_{t-1}^{ib\ell}\right).$$

But, the opposite pattern arises if we instead use the first trade **a**fter time ℓ,

$$\text{cov}\left(\Delta p_t^{ia\ell}, \Delta p_{t-1}^{fa\ell}\right) = (1-\rho_f)\rho_i N\text{var}(\delta_t^{\ell+1}) < \rho_f(1-\rho_i)N\text{var}(\delta_t^{\ell+1}) = \text{cov}\left(\Delta p_t^{fa\ell}, \Delta p_{t-1}^{ia\ell}\right).$$

Thus, the qualitative insights from the two stock example extend when we consider portfolios.

Delayed Incorporation of Information into Prices. We next derive how delays in the incorporation of information into prices affect autocorrelation patterns. There is both evidence that (i) information arrival drives trading frequency, so that a stock that trades more frequently than usual is likely to reflect information arrival (see Hollifield, et al. (2003)), and (ii) transaction prices of infrequently-traded stocks may reflect lagged valuations. These observations are crucially different: the first is consistent with efficient markets, while the second is not.

If trade frequency is driven only by information arrival, then the impact of nonsynchronous trading on correlation patterns is unaltered: in our analysis, this simply provides an interpretation of which stocks in a given portfolio happen to trade frequently, i.e., to trade close to time ℓ. But what happens if information is incorporated into some stocks with a *lag*?

To present this most simply, suppose that all innovations have the same variance, σ^2, and that a stock price either reflects all relevant information, or the price reflects information with a lag of $z > 0$. Let $\pi_f(z)$ and $\pi_i(z)$ be the respective fraction of frequently- and infrequently-traded stocks that have a lag of z, where we assume that z is small enough that the stock price reflects the stock's value sometime within the previous 24 hours.

Then the covariance between the lagged returns of the frequently-traded portfolio and the current returns of the infrequently-traded portfolio before time ℓ is

$$\begin{aligned}\mathrm{cov}(\Delta p_{t-1}^{fb\ell}, \Delta p_t^{ib\ell}) &= N\sigma^2\Big[(\rho_f\rho_i + (1-\rho_f)(1-\rho_i))\pi_f(0)\pi_i(z)z + \rho_f(1-\rho_i)[\pi_f(0)\pi_i(0)(1) \\ &\quad +\pi_f(0)\pi_i(z)(z+1) + \pi_f(z)\pi_i(z)(1)] + (1-\rho_f)\rho_i\pi_f(0)\pi_i(z)(z-1)\Big],\end{aligned}$$

and the corresponding covariance after time ℓ is

$$\begin{aligned}\mathrm{cov}(\Delta p_{t-1}^{fa\ell}, \Delta p_t^{ia\ell}) &= N\sigma^2\Big[(\rho_f\rho_i + (1-\rho_f)(1-\rho_i))\pi_f(0)\pi_i(z)z + \rho_i(1-\rho_f)[\pi_f(0)\pi_i(0)(1) \\ &\quad +\pi_f(0)\pi_i(z)(z+1) + \pi_f(z)\pi_i(z)(1)] + (1-\rho_i)\rho_f\pi_f(0)\pi_i(z)(z-1)\Big].\end{aligned}$$

Taking the difference between these two covariances yields

$$\begin{aligned}
& N\sigma^2(\rho_f-\rho_i)\left[\pi_i(0)\pi_f(0)(1)+\pi_i(0)\pi_i(z)(z+1)+\pi_f(z)\pi_i(z)(1)-\pi_f(0)\pi_i(z)(z-1)\right] \\
= \;& N\sigma^2(\rho_f-\rho_i)\left[\pi_i(0)\pi_f(0)+2\pi_i(0)\pi_i(z)+\pi_f(z)\pi_i(z)\right] \\
= \;& N\sigma^2(\rho_f-\rho_i)\left(1+\pi_f(0)-\pi_i(0)\right).
\end{aligned}$$

If a greater fraction of infrequently-traded stock prices reflect lagged information (relative to frequently-traded stock prices) as the day progresses, then, all else equal, cross-autocorrelations should rise throughout the day as should before-after cross-autocorrelations differences.

Further, to the extent that frequently-traded stocks and infrequently-traded stocks are composed of stocks that systematically vary in the extent to which they trade and/or incorporation information, then before-after patterns similar to those found with cross-autocorrelations emerge with autocorrelations. To illustrate, suppose for the fraction ρ_f of frequently-traded stocks that trade "within an innovation" of time ℓ, fraction $\pi_{ff}(z)$ have prices that reflect lagged information; and for the fraction $1-\rho_f$ of frequently-traded stocks that trade more than an innovation of time ℓ, fraction $\pi_{fi}(z)$ have prices that reflect lagged information, where $\pi_{fi}(z) > \pi_{ff}(z)$. Intuitively, this captures that within the portfolio of stocks that tend to trade frequently, those that locally traded near a moment of time ℓ are more likely to be traded frequently at that moment, and the fact that they are more frequently-traded means that substantial information arrival is more likely, information that is more likely to be incorporated into price. In contrast, the cost of not keeping up with a stock with less information arrival is less; such stocks trade less frequently as a result, and their prices are more likely not to reflect all extant information.

Then the own-autocovariance for the portfolio of frequently-traded stocks before time ℓ is:

$$\mathrm{cov}(\Delta p_{t-1}^{fb\ell},\Delta p_t^{fb\ell}) = N\sigma^2[(1-\rho_f)\pi_{fi}(z)[\rho_f\pi_{ff}(z)+(1-\rho_f)\pi_{fi}(0)z+\rho\pi_{ff}(0)(z+1)]$$

$$+\rho_f\pi_{ff}(z)[(1-\rho_f)\pi_{fi}(0)(z-1)+\rho_f\pi_{ff}(0)z]+(1-\rho_f)\rho_f\pi_{fi}(0)\pi_{ff}(0)]$$

and the own-autocovariance after time ℓ is:

$$\text{cov}(\Delta p_{t-1}^{fa\ell}, \Delta p_t^{fa\ell}) = N\sigma^2[\rho_f\pi_{ff}(z)[(1-\rho_f)\pi_{fi}(z) + \rho_f\pi_{ff}(0)z + (1-\rho_f)\pi_{fi}(0)(z+1)]$$

$$+(1-\rho_f)\pi_{fi}(z)[\rho_f\pi_{ff}(0)(z-1) + (1-\rho_f)\pi_{fi}(0)z] + \rho_f\pi_{ff}(0)(1-\rho_f)\pi_{fi}(0)]$$

The difference between before and after own-autocovariances is:

$$\text{cov}(\Delta p_{t-1}^{fb\ell}, \Delta p_t^{fb\ell}) - \text{cov}(\Delta p_{t-1}^{fa\ell}, \Delta p_t^{fa\ell}) = 2N\sigma^2\rho_f(1-\rho_f)[\pi_{ff}(0) - \pi_{fi}(0)]$$

If $\pi_{ff}(0) \sim \pi_{fi}(0)$, then there should be little or no difference between the before and after own-autocovariances. If $|\pi_{ff}(0)-\pi_{fi}(0)| < |\pi_{if}(0)-\pi_{ii}(0)|$, then the difference in before/after own-autocovariances should be larger for portfolios of infrequently-traded stocks than for portfolios of frequently-traded stocks.

4 Data

Trade and quote data are obtained from the TAQ database for the period 1993 to 1998. From Datastream, we obtain the market capitalization on January 1st of each year for all NYSE-listed stocks.[4] We eliminate stocks for which Datastream does not have market capitalization data, stocks based outside the United States, and securities that are not ordinary common shares. Finally, we eliminate stocks that do not trade at least 200 times each month and stocks that change their ticker symbol during the year. From the remaining sample, on a yearly basis we select the largest 500 firms by market capitalization and sort these into two portfolios: (i) the largest (L) 250 stocks and (ii) the smallest (S) 250 stocks. The same 500 firms are also divided into two groups according to their trade volume in the first month of the year: (i) 250 frequently-traded (f) stocks and (ii) 250 infrequently-traded (i) stocks.[5]

[4] Results for Nasdaq-listed stocks were found to be qualitatively similar and are not presented here to improve clarity and length. They are available from the authors upon request.

[5] Barberis and Shleifer (2003) find that assets grouped according to a "style" exhibit even stronger own- and cross-autocorrelation patterns at a *monthly* return horizon. To the extent that style-based flows change on a daily horizon, our results might be different if we divide the stocks according to style-based characteristics.

We focus on actively-traded, large stocks it is easiest to interpret our results when stocks trade in (virtually) all 24-hour periods. Later, we consider how our results might change if we include highly-illiquid stocks.

We consider $\tau \in \{\text{market open}, 1100h, 1200h, 1245h, 1330h, 1430h, \text{market close}\}$ as arbitrary times. We focus on these times for our statistical analysis because they are sufficiently far apart to keep overlap in trade times to a minimum. For each stock, a series of daily prices are created for each of the following criteria: first trade price after time τ, last trade price before time τ, the mid-quote of the first quote revision after time τ, and the mid-quote of the last quote revision prior to time τ. Based on this series of prices, daily portfolio returns are calculated as: $r_t = \frac{1}{I}\sum_{i=1}^{I}\frac{P_{ti}-P_{t-1,i}}{P_{t-1,i}}$. Note that the last trade prior to an arbitrary time might occur on the previous trading day and the first trade after an arbitrary time might occur on the next trading day (within 24 hours). We eliminate potentially erroneous prices for which the absolute daily price change is greater than 50% or for which the bid price exceeds the ask price. In the absence of suitable prices, we assume that the daily return for that particular stock is zero.

In our sample, on average, more than 99.8% of stocks have a trade/quote during each 24-hour time period over which returns are calculated—thus, non-trading over multiple days is not a significant concern. Non-trading over multiple days reduces cross-autocorrelation patterns in daily returns, because there is then no overlap in information arrival.

5 Results

5.1 Empirical Cumulative Distribution of Transaction Times

Figure 3 illustrates the empirical cumulative distribution of times of the first trade after 0930h for the large capitalization and small capitalization stock portfolios. The trade times of large capitalization firms stochastically dominate (first-order) the trade times of small capitalization firms. Note that there is surprisingly little difference in the distribution of opening trade times between the large and small stocks samples.

Figure 4 illustrates the analogous empirical cumulative distribution of the time of the last trade prior to 1600h. Previous studies of cross-autocorrelation patterns are based on official closing prices, which may reflect market-on-close (MOC) and limit-on-close (LOC) trades. MOC and LOC orders must be submitted prior to 1540h.[6] When the market closes at 1600h, all MOC orders are executed against the prevailing bid or offer. An imbalance of buy orders would be executed against the offer side of the market and an imbalance of sell orders, the bid side. If an imbalance still remains at 1600h, the specialist executes a proprietary trade to alleviate the disparity between supply and demand or seek approval for a trading halt if a significant order imbalance remains after publication and receipt of any offsetting orders. In general, most of the delay in reporting times after 1600h in figure 5 reflect the arbitrary time the specialist and his clerk decide to report the trade, rather than information-driven factors. Note that the stochastic dominance in closing trade times is still present when MOC and LOC trades are considered.

In addition to the open and close, we conduct our before/after autocorrelation analysis at different times during the day. Figure 6 provides the distribution of trade times around 1245h. We highlight this trading time because it divides the trading day in half. This time is sufficiently far away from the open and close that we can avoid problems associated with the opening call auction and different rates of information release at the open and close, thereby examining the nonsynchronous trading effect "directly". Our analysis is most transparent when trade times are symmetrically distributed around our arbitrary times, as figure 6 verifies.

Figures 7 and 8 provide the corresponding empirical cumulative distributions of the time of the first and last quote revision, respectively. It is worth noting that there is little difference in the distribution of first quote times for the large and small stock portfolios.

[6] All MOC and LOC orders must be entered by 1540h either systemically or by floor brokers representing MOC and LOC orders who must communicate their interest to the specialist trading the stock. After 1540h, MOC and LOC orders can only be entered if they are on the contra side of the last imbalance published.

5.2 Autocorrelation Patterns

Table 1 presents the cross-autocorrelation results for size-sorted (large and small) portfolios with returns calculated using both trade prices and quote midpoints. The distribution of autocorrelation estimates is highly non-normal and exhibits significant skew. Thus, using asymptotic standard errors to construct *symmetric* confidence intervals for our autocorrelation estimates will produce confidence bounds that are too wide and incorrectly centered around our estimate. Since the difference in autocorrelation estimates measured before and after a particular point in time is potentially very small, it is important that we construct confidence intervals that reflect the underlying true distribution as closely as possible. To do so, we construct bootstrap confidence intervals for the autocorrelation estimates.

First, we observe that using a standard block bootstrap approach to estimate autocorrelations is inappropriate because a substantial proportion of the pairs in a resampled series will lie across a join between blocks, and will therefore be independent. To overcome this whitening effect of block resampling, we instead use a blocks of blocks bootstrap approach designed to preserve the time dependence structure in the original data. See Davison and Hinkley (1997, p. 398) for details. Suppose our original data series is $(y_1, y_2, \ldots, y_n)$ and we seek to calculate a lag 1 autocovariance. Then, we set

$$(y'_1, \ldots, y'_{n-1}) = \begin{pmatrix} y'_{11} & y'_{12} & \cdots & y'_{1,n-1} \\ y'_{21} & y'_{22} & \cdots & y'_{2,n-1} \end{pmatrix} = \begin{pmatrix} y_1 & y_2 & \cdots & y_{n-1} \\ y_2 & y_3 & \cdots & y_n \end{pmatrix}.$$

We then resample blocks of the new "data" $y'_1, \ldots, y'_{n-m+1}$, each of the observations of which is a block of the original data. The key point is that our statistic of interest should not compare observations adjacent in each row. With $n = 12$ observations and a block length of $l = 4$ a bootstrap replicate might be

$$\{y'^*_j\} = \begin{pmatrix} y_5 & y_6 & y_7 & y_8 & & y_1 & y_2 & y_3 & y_4 & & y_7 & y_8 & y_9 & y_{10} \\ y_6 & y_7 & y_8 & y_9 & & y_2 & y_3 & y_4 & y_5 & & y_8 & y_9 & y_{10} & y_{11} \end{pmatrix}.$$

For each bootstrap replication $b = 1, \ldots, B$, the statistic of interest, t_b, is calculated based on the corresponding bootstrap replicate. The number of replications B is selected such that $\alpha(B + 1)$ is an integer. A basic (percentile) bootstrap confidence interval with nominal coverage of $(1-\alpha)$ is obtained by sorting the statistics t_b from largest to smallest, $t_1^*, t_2^*, \ldots, t_B^*$ and then constructing the confidence interval with confidence limits $\hat{\theta}_\alpha = 2t - t^*_{(B+1)(1-\alpha)}$ and $\hat{\theta}_{1-\alpha} = 2t - t^*_{(B+1)\alpha}$.

We construct confidence intervals with $B = 9999$ bootstrap replications and using a block length of $l = 5$. Robustness checks show that our results are not sensitive to using different "reasonable" block lengths. Table 2 presents the confidence intervals with 95% nominal coverage obtained using this blocks of blocks bootstrap technique. Our estimated confidence intervals for the reported autocorrelations have widths of approximately 0.14. That is, it is difficult to establish the *level* of correlation precisely. In contrast, we can compare two correlations by taking their difference and then constructing a bootstrap confidence interval of the difference. The confidence intervals for the difference have widths of about 0.01 and thus provide a *very* precise estimate of the relative orderings of the correlations. For instance, we show that the confidence interval of the difference $\text{corr}(r_{t-1}^{Lb}, r_t^{Sb}) - \text{corr}(r_{t-1}^{La}, r_t^{Sa})$ bounds strictly positive numbers, while the confidence interval of the difference $\text{corr}(r_{t-1}^{Sb}, r_t^{Lb}) - \text{corr}(r_{t-1}^{Sa}, r_t^{La})$ bounds strictly negative numbers.

Consistent with previous research, we find that when closing prices are used to calculate returns, $\text{corr}(r_{t-1}^{L}, r_t^{S})$ is significantly positive and far exceeds $\text{corr}(r_{t-1}^{S}, r_t^{L})$. In fact, regardless of when we compute returns, we find that $\text{corr}(r_{t-1}^{L}, r_t^{S}) >> \text{corr}(r_{t-1}^{S}, r_t^{L})$, i.e., the correlation of lagged returns of large-capitalization stocks with current returns of small-capitalization stocks always exceeds the "reverse" cross-autocorrelation. This can be easily verified by constructing confidence intervals of the difference between them using the blocks of blocks bootstrap approach.

We find clear evidence that the impact of nonsynchronous trading is very small. For example, the predictive power of large stocks is slightly higher using the last trade before 1245h, $\text{corr}(r_{t-1}^{Lb}, r_t^{Sb}) = 0.104$, than using the first trade after 1245h, $\text{corr}(r_{t-1}^{La}, r_t^{Sa}) = 0.101$; while the predictive power of small stocks is slightly higher using the first trade after 1245h, $\text{corr}(r_{t-1}^{Sa}, r_t^{La}) = 0.078$, than using last trade before 1245h, $\text{corr}(r_{t-1}^{Sb}, r_t^{Lb}) = 0.073$. In fact, the corresponding confidence intervals of the difference in correlations $[-0.002, 0.008]$ and

$[-0.001, 0.010]$ both include zero. Thus, the correlations at 1245h move in exactly the direction predicted by the impact of nonsynchronous trading, but the impact is slight (and often insignificant from zero).

Further evidence of the slight impact of nonsynchronous trading can be gleaned from the before-after own-autocorrelation patterns. Specifically, it has been well-established that information arrival drives trading frequency. As a result, within a portfolio stocks that trade frequently will tend to have more information. As we have seen using the last trade before a time ℓ creates an time overlap for lagged frequently-traded stocks with current infrequently-traded stocks, while using the first trade after ℓ creates the opposite time overlap. As a result, the own-autocorrelation pattern should mirror the cross-serial correlation pattern of lagged frequently-traded portfolio returns with current infrequently-traded portfolio returns. Consistent with this, table 1 reveals that own portfolio lagged autocorrelation is slightly higher using the last trade before than using the first trade after. Again, this slight difference indicates that nonsynchronous trading has only a minor impact.

The fact that $\text{corr}(r_{t-1}^{La}, r_t^{Sa})$ is substantially greater than $\text{corr}(r_{t-1}^{Sa}, r_t^{La})$ strongly indicates that pricing of infrequently-traded stocks reflect some inefficiency—when prices of infrequently-traded stocks are determined, they must not incorporate some of the recent information that is already contained in the prices of frequently-traded stocks. This is because the impact of nonsynchronous trading in the absence of delayed incorporation of information would give rise to the opposite inequality.

The own-autocorrelation patterns provide added evidence that recent information is not immediately incorporated into prices of infrequently-traded stocks. No matter whether we use the last trade before some time ℓ, or the first trade after ℓ, $\text{corr}(r_{t-1}^{S}, r_t^{S})$ is always large and positive (exceeding the cross-portfolio lead-lag autocorrelations), and greatly exceeding $\text{corr}(r_{t-1}^{L}, r_t^{L})$. This can be reconciled if information is impounded more quickly into some stocks than others, especially for small capitalization stocks.

Our analysis of autocorrelation patterns at different points in the trading day uncovers several other strong empirical regularities. First, the before-after difference in cross-autocorrelations is very consistent throughout the trading day. Specifically at 1100h, 1200h, 1245h, 1330h, and 1430h, we find that $\text{corr}(r_{t-1}^{Lb}, r_t^{Sb}) > \text{corr}(r_{t-1}^{La}, r_t^{Sa})$, while $\text{corr}(r_{t-1}^{Sb}, r_t^{Lb}) < \text{corr}(r_{t-1}^{Sa}, r_t^{La})$. This result is robust: it holds no matter whether we compute returns using trades or quotes. Nonetheless, the impact of nonsynchronous trading remains negligible at all points in the trading day. We also find quite generally that the autocorrelation patterns are very similar whether we compute returns using trades or quotes: the extent to which the bid-ask bounce in trade-to-trade returns induces first-order autocorrelation does not underlie any of our findings.

Second, autocorrelations are increasing throughout the trading day, prior to decreasing slightly at the market close. This effect is dramatic. For example, from 1100 to close, $\text{corr}(r_{t-1}^{L}, r_t^{S})$, essentially triples, while the own correlation of small-capitalization stocks $\text{corr}(r_{t-1}^{S}, r_t^{S})$ doubles.

In order to confirm that the statistical significance of the increasing autocorrelations during the day, for each of 9999 bootstrap replications, using the block of blocks bootstrap method with a block length of 40,[7] we perform a nonparametric Wilcoxon Signed-Rank test based on the following 16 differences:

$$
\begin{array}{ll}
\text{corr}(r_{t-1}^{L}, r_t^{L})_{time=1430} - \text{corr}(r_{t-1}^{L}, r_t^{L})_{time=1245} & \text{corr}(r_{t-1}^{S}, r_t^{S})_{time=1430} - \text{corr}(r_{t-1}^{S}, r_t^{S})_{time=1245} \\
\text{corr}(r_{t-1}^{L}, r_t^{L})_{time=1245} - \text{corr}(r_{t-1}^{L}, r_t^{L})_{time=1100} & \text{corr}(r_{t-1}^{S}, r_t^{S})_{time=1245} - \text{corr}(r_{t-1}^{S}, r_t^{S})_{time=1100} \\
\text{corr}(r_{t-1}^{L}, r_t^{L})_{time=1430} - \text{corr}(r_{t-1}^{L}, r_t^{L})_{time=1100} & \text{corr}(r_{t-1}^{S}, r_t^{S})_{time=1430} - \text{corr}(r_{t-1}^{S}, r_t^{S})_{time=1100} \\
\text{corr}(r_{t-1}^{L}, r_t^{L})_{time=1330} - \text{corr}(r_{t-1}^{L}, r_t^{L})_{time=1200} & \text{corr}(r_{t-1}^{S}, r_t^{S})_{time=1330} - \text{corr}(r_{t-1}^{S}, r_t^{S})_{time=1200} \\
\\
\text{corr}(r_{t-1}^{L}, r_t^{S})_{time=1430} - \text{corr}(r_{t-1}^{L}, r_t^{S})_{time=1245} & \text{corr}(r_{t-1}^{S}, r_t^{L})_{time=1430} - \text{corr}(r_{t-1}^{S}, r_t^{L})_{time=1245} \\
\text{corr}(r_{t-1}^{L}, r_t^{S})_{time=1245} - \text{corr}(r_{t-1}^{L}, r_t^{S})_{time=1100} & \text{corr}(r_{t-1}^{S}, r_t^{L})_{time=1245} - \text{corr}(r_{t-1}^{S}, r_t^{L})_{time=1100} \\
\text{corr}(r_{t-1}^{L}, r_t^{S})_{time=1430} - \text{corr}(r_{t-1}^{L}, r_t^{S})_{time=1100} & \text{corr}(r_{t-1}^{S}, r_t^{L})_{time=1430} - \text{corr}(r_{t-1}^{S}, r_t^{L})_{time=1100} \\
\text{corr}(r_{t-1}^{L}, r_t^{S})_{time=1330} - \text{corr}(r_{t-1}^{L}, r_t^{S})_{time=1200} & \text{corr}(r_{t-1}^{S}, r_t^{L})_{time=1330} - \text{corr}(r_{t-1}^{S}, r_t^{L})_{time=1200}
\end{array}
\tag{2}
$$

For portfolios sorted based on market capitalization, the test statistic is significant at the 5% level for 9176 replications when returns are based on trade prices, implying a bootstrap p-value of 0.0824. For quotes, the bootstrap p-value is 0.0860. For portfolios sorted based

[7] A longer block length is used to place greater emphasis on the *relative ordering* of the correlations rather than the *level* of the correlations.

on the value of trading volume, the bootstrap p-value is 0.0689 for trades and 0.0617 for quotes. Thus, we conclude that the intraday pattern of rising autocorrelations is significant.

The intraday pattern can be explained as follows. During the overnight period, investors gather information and stock prices "catch up" with their true underlying valuations. During the day, the incorporation of new information into stock prices becomes increasingly lagged, particularly for small stocks. Then, the differences at the close are consistent with large stocks getting "fresher" at the close. Autocorrelation patterns at the close are not representative of the patterns throughout the remainder of the day. The close magnifies the power of the large stock portfolio returns to predict future small stock portfolio returns. It is important to emphasize that this is due to the delayed incorporation of information, and not due to lower trading frequencies later in the day.

Samples based on trade frequency: We now repeat the analysis by constructing samples based on trade frequency. Using the same set of stocks used in the market capitalization analysis, we re-sort according to the number of trades in the first month of the year. The sample is then divided into frequently-traded and infrequently-traded portfolios of 250 stocks. Despite only a loose correspondence between market capitalization and trade frequency (see figure 9), the results in table 3 are qualitatively similar to those found previously. As expected, the nonsynchronous trading effect is larger when we sort on trading frequency since there is a greater variation in trade times between the frequently and infrequently portfolios. To verify this, table 4 reports the corresponding confidence intervals using the blocks of blocks bootstrap approach: many of the confidence intervals that previously straddled zero for size sorted portfolios are now strictly positive for trade frequency sorted portfolios.

6 Conclusion

This paper re-examines the extent to which differences in trading frequencies can underlie lead-lag cross-autocorrelations in stock returns. Our central insight is that by using trade prices before and after an arbitrary point in the trading day to calculate returns, we can isolate the impact of nonsynchronous trading on portfolio correlation patterns. This permits us to conclude decisively that the impact of nonsynchronous trading is negligible. Rather, the autocorrelation patterns reflect inefficient pricing of infrequently-traded stocks—when prices of infrequently-traded stocks are determined, they do not incorporate some of the recent information that is already contained in the prices of frequently-traded stocks.

Our analysis suggests several directions for future research. First, more detailed order submission information could be used to establish how much "staleness" is due to infrequently updated limit orders compared with other sources of staleness. Second, the time series could be extended in order to examine how intraday autocorrelation patterns have changed over time as institutional features (e.g. tick size) and trading frequency have changed.

References

Ahn, D.H., J. Boudoukh, M. Richardson, R.F. Whitelaw, 2002, Partial adjustment or stale prices? Implications from stock index and futures return autocorrelations, Review of Financial Studies 15(2): 655–689.

Atchison, M., K. Butler, R. Simonds, 1987, Nonsynchronous security trading and market index autocorrelation, Journal of Finance 42(1): 111–118.

Badrinath, S.G., J.R. Kale, T.H. Noe, 1995, Of shepherds, sheep, and the cross-autocorrelations in equity returns, Review of Financial Studies 8(2): 401–430.

Barberis, N., A. Shleifer, 2003, Style investing, Journal of Financial Economics 68: 161–199.

Bernhardt, D., E. Hughson, 2002, Intraday Trade in Dealership Markets, European Economic Review 46(6): 1697–1732.

Bessembinder, H., M. Hertzel, 1993, Return autocorrelation around nontrading days, Review of Financial Studies 6: 155–189.

Boudoukh, J., M. Richardson, R. Whitelaw, 1994, A tale of three schools: Insights on autocorrelation of short-horizon returns, Review of Financial Studies 7(3): 539–573.

Brennan, M.J., N. Jagadeesh, B. Swaminathan, 1993, Investment Analysis and The Adjustment of Stock Prices to Common Information, Review of Financial Studies 6: 799–824.

Campbell, J.Y., Grossman, S.J., Wang, J., 1993, Trading volume and serial correlation in stock returns. Quarterly Journal of Economics 108(4): 905–939.

Chan, K., 1992, A further analysis of the lead-lag relationship between the cash market and stock index futures markets, Review of Financial Studies 5(1): 123–152.

Chan, K., 1993, Imperfect information and cross-autocorrelation among stock prices, Journal of Finance 48(4): 1211–1230.

Chalmers, J.M.R., R.M. Edelen, G.B. Kadlec, 2001, On the perils of financial intermediaries setting security prices: The mutual fund wild card option, Journal of Finance 56(6): 2209–2236.

Chordia, T., B. Swaminathan, 2000, Trading volume and cross-autocorrelations in stock returns, Journal of Finance 55: 913–936.

Cohen, K.J., S.F. Maier, R.A. Schwartz, D.K. Whitcomb, 1986, The microstructure of securities markets (Prentice-Hall, Englewood Cliffs, NJ).

Conrad, J., G. Kaul, 1988, Time varying expected returns, Journal of Business 61: 409–425.

Conrad, J., G. Kaul, 1989, Mean-reversion in short-horizon individual security return, Journal of Financial Economics 29: 365–384.

Cornell, B., 1985, The Weekly Pattern in Stock Returns: Cash versus Futures, Journal of Finance, 40: 583–588.

Davison, A.C., D.V. Hinkley, 1997, Bootstrap methods and their application, Cambridge University Press.

Fisher, L., 1966, Some new stock market indexes, Journal of Business 39(1): 191–225.

Hameed, A., 1997, Time-varying factors and cross-autocorrelations in short-horizon stock returns, Journal of Financial Research 20: 435–58.

Hameed, A., Y. Kusnadi, 2003, Stock return cross-autocorrelations and market conditions in Japan, Working Paper, National University of Singapore.

Hollifield, B., R.A. Miller, P. Sandas, J. Slive, 2003, Liquidity supply and demand in limit order markets, Working Paper, Carnegie Mellon University.

Hou, K., 2003, Industry information diffusion and the lead-lag effect in stock returns, Working Paper, Ohio State University.

Kadlec, G.B., D.M. Patterson, 1999, A transactions data analysis of nonsynchronous trading, Review of Financial Studies 12(3): 609–630.

Lewellen, J., 2002, Momentum profits and the autocorrelation of stock returns, Review of Financial Studies 15(2): 533–563.

Lo, A., A.C. MacKinlay, 1990, When are contrarian profits due to stock market overreaction?, Review of Financial Studies 3(2): 175–205.

Lo, A., A.C. MacKinlay, 1990, An econometric analysis of nonsynchronous trading, Journal of Econometrics, 45: 181–211.

McInish, T.H, R.A. Wood, 1991, Autocorrelation of Daily Index Returns: Intraday-to-Intraday Versus Close-to-Close Intervals, Journal of Banking and Finance, 15(1): 193–206.

McQueen, G., J.M. Pinegar, S. Thorley, 1996, Delayed reaction to good news and the cross-autocorrelation of portfolio returns, Journal of Finance 51(3), 889–919.

Mech, T., 1993, Portfolio return autocorrelation, Journal of Financial Economics 34: 307–44.

Olson, D., C. Mossman, 2001, Cross-correlations and predictability of stock returns, Journal of Forecasting 20(2): 145–160.

Perry, P.R., 1985, Portfolio serial correlation and nonsynchronous trading, Journal of Financial and Quantitative Analysis 20: 517–523.

Richardson, M., T. Smith, 1994, A unified approach to testing for serial correlation in stock returns, Journal of Business 67(3): 371–399.

Richardson, T., D.R. Peterson, 1999, The cross-autocorrelation of size-based portfolio returns is not an artifact of portfolio autocorrelation, Journal of Financial Research, 12(1):

1–13.

Roll, R., 1984, A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market, Journal of Finance 39(4): 1127–1139.

Scholes, M., J. Williams, 1977, Estimating betas from nonsynchronous data, Journal of Financial Economics 5(3): 309–327.

Sentana, E., Wadhwani, S., 1992, Feedback traders and stock return autocorrelation: Evidence from a century of daily data, Economic Journal 102: 415–425.

Schwartz, R.A., D.K. Whitcomb, 1977, The time-variance relationship: Evidence on autocorrelation in common stock returns, Journal of Finance 32: 41–55.

Sias, R.W., L.T. Starks, 1997, Return autocorrelation and institutional investors, Journal of Financial Economics 46: 103–131.

Stoll, H.R., R.E. Whaley, 1990, The dynamics of stock index and stock index futures returns, Journal of Financial and Quantitative Analysis 25(4): 441–468.

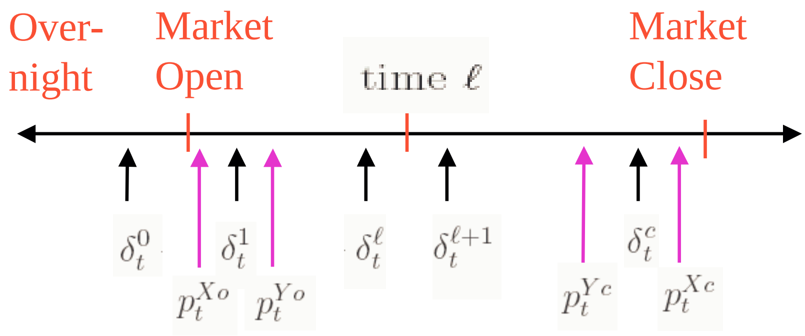

Figure 1: **Timeline.** Value of each stock at end of date t is given by $V_t = V_{t-1} + \delta_t^0 + \delta_t^1 + \ldots + \delta_t^\ell + \delta_t^{\ell+1} + \ldots + \delta_t^c$, where the innovations δ_t^j are independently distributed. Before and after each innovation, stock X is traded at least once. In contrast, first trade of stock Y always occurs just after δ_t^1 has been realized; and Y last trades just before δ_t^c has been realized.

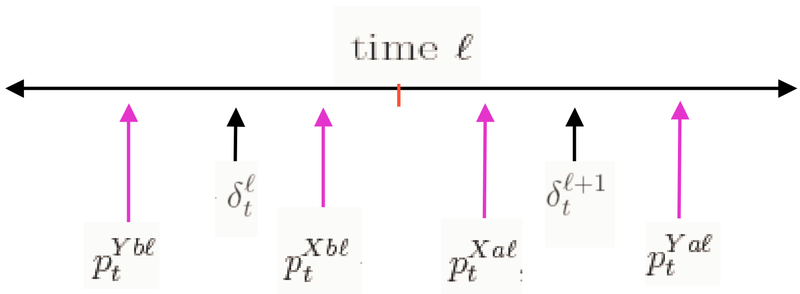

Figure 2: **Arbitrary point of time.** The innovation δ_t^ℓ occurs before time ℓ and is incorporated into the time ℓ price of stock X, but not of stock Y; The first trade after ℓ in stock X occurs before $\delta_t^{\ell+1}$ is realized, but that the first trade in Y occurs after.

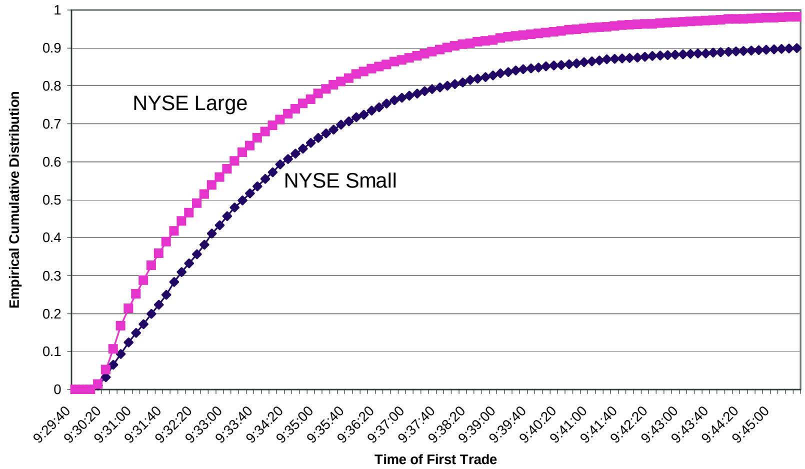

Figure 3: The empirical cumulative distribution at 10 second intervals of the time elapsed from 0930h to the first trade. The sample period is January 1993. The empirical cumulative distribution of opening trade times is $\hat{F}(\tau) = (1/250) \times \sum_{i=1}^{250} I(t_i^o < \tau)$, where $I(\cdot)$ is an indicator function that equals 1 if true, zero otherwise; t_i^o is the opening trade time of stock i.

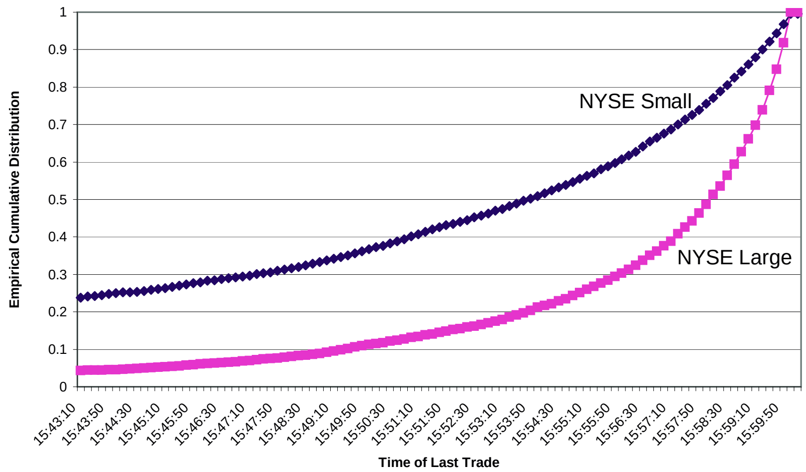

Figure 4: The empirical cumulative distribution at 10 second intervals of the time of the last trade prior to 1600h. The sample period is January 1993. The empirical cumulative distribution of closing trade times is $\hat{F}(\tau) = (1/250) \times \sum_{i=1}^{250} I(t_i^c < \tau)$, where $I(\cdot)$ is an indicator function that equals 1 if true, zero otherwise; t_i^c is the closing trade time of stock i.

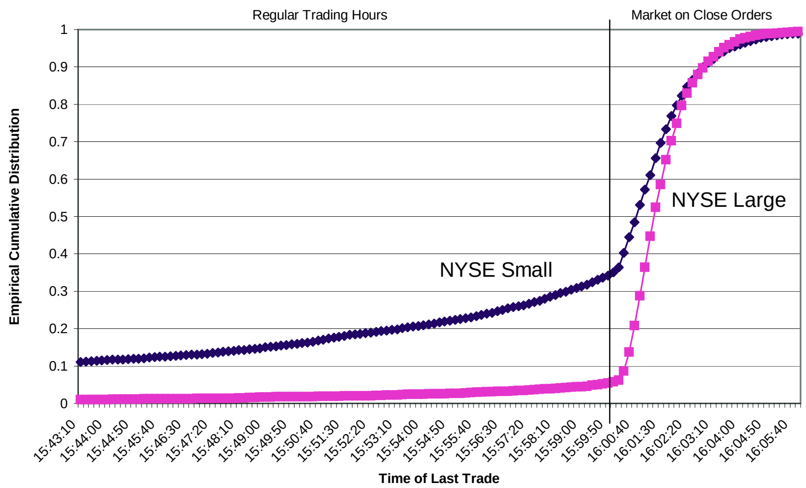

Figure 5: The empirical cumulative distribution at 10 second intervals of the time of the official last trade, including market-on-close and limit-on-close orders. The sample period is January 1993. The empirical cumulative distribution of closing trade times is $\hat{F}(\tau) = (1/250) \times \sum_{i=1}^{250} I(t_i^c < \tau)$, where $I(\cdot)$ is an indicator function that equals 1 if true, zero otherwise; t_i^c is the closing trade time of stock i.

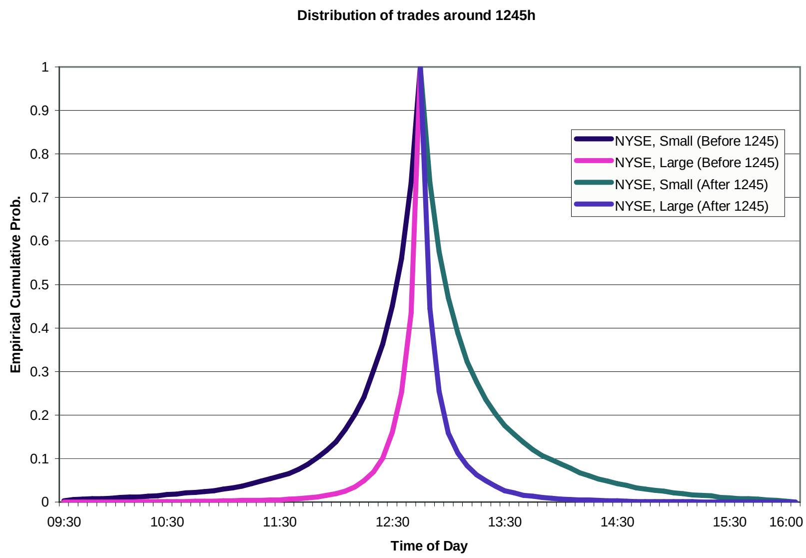

Figure 6: The empirical cumulative distribution of the time elapsed from the last trade prior to 1245h and after 1245h. The sample period is January 1993. The empirical cumulative distribution of trade times prior to 1245h is $\hat{F}(\tau) = (1/250) \times \sum_{i=1}^{250} I(t_i^{b1245} < \tau)$, where $I(\cdot)$ is an indicator function that equals 1 if true, zero otherwise; t_i^{b1245} is the time of the last trade of stock i prior to 1245. The empirical cumulative distribution of trade times after 1245h is $\hat{F}(\tau) = (1/250) \times \sum_{i=1}^{250} I(t_i^{a1245} < \tau)$, where $I(\cdot)$ is an indicator function that equals 1 if true, zero otherwise; t_i^{a1245} is the time of the last trade of stock i after 1245.

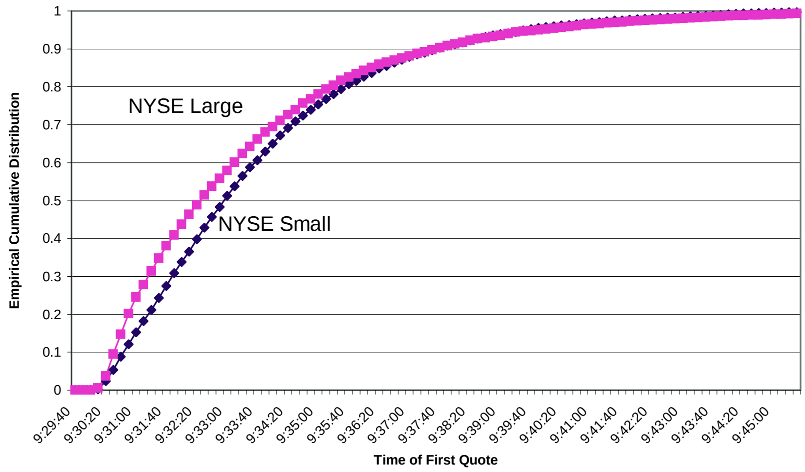

Figure 7: The empirical cumulative distribution at 10 second intervals of the time elapsed from 0930h to the first quote. The sample period is January 1993. The empirical cumulative distribution of opening quote revision times is $\hat{F}(\tau) = (1/250) \times \sum_{i=1}^{250} I(t_i^o < \tau)$, where $I(\cdot)$ is an indicator function that equals 1 if true, zero otherwise; t_i^o is the time of the first quote revision of stock i.

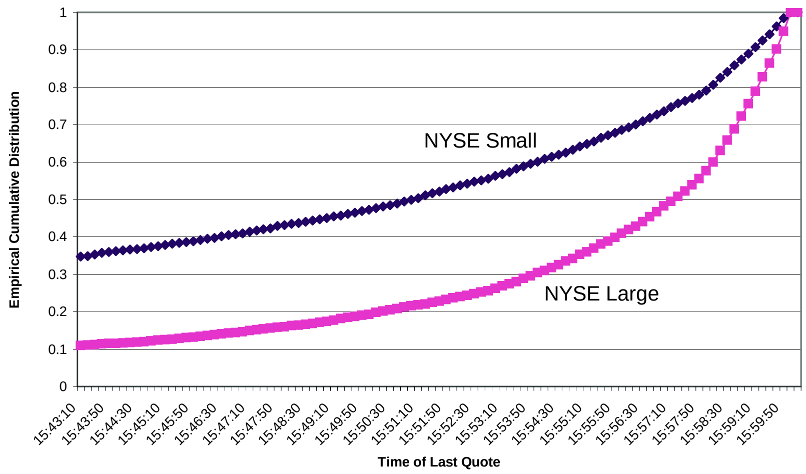

Figure 8: The empirical cumulative distribution at 10 second intervals of times of last quote prior to 1600h. The sample period is January 1993. The empirical cumulative distribution of closing quote revision times is $\hat{F}(\tau) = (1/250) \times \sum_{i=1}^{250} I(t_i^c < \tau)$, where $I(\cdot)$ is an indicator function that equals 1 if true, zero otherwise; t_i^c is the time of the last quote revision of stock i.

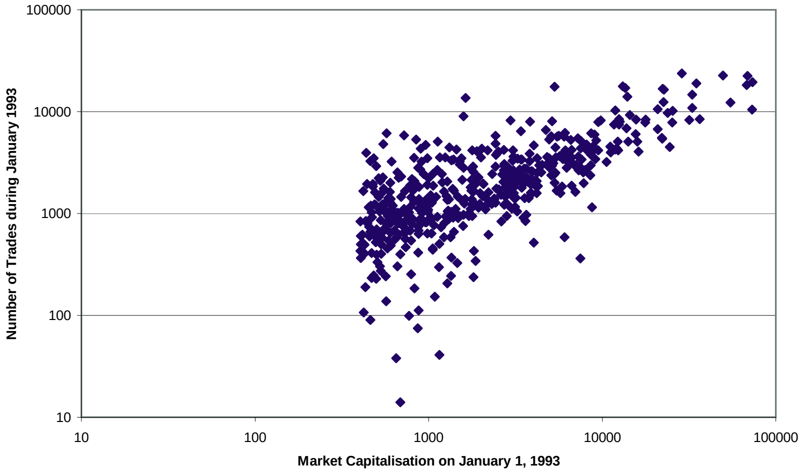

Figure 9: Scatter plot of market capitalization on January 1, 1993 versus the number of trades during January 1993 for the sample of 500 stocks. Logarithmic scale is used on both axes.

Table 1: **Autocorrelation patterns of NYSE portfolios sorted by market capitalization** (obtained on annual basis from Datastream). Only common shares of companies whose primary geographical location is the US are considered. The large (L) and the small (S) market capitalization portfolios each have 250 stocks. Sample period is 1993–1998. $<\tau$h indicates the last trade prior to τh. $>\tau$h indicates the first trade after τh.

Trades

	Open	<1100h	>1100h	<1200h	>1200h	<1245h	>1245h	<1330h	>1330h	<1430h	>1430h	Close
$\text{corr}(r_t^L, r_t^S)$	0.908	0.924	0.926	0.917	0.919	0.922	0.924	0.922	0.923	0.918	0.917	0.900
$\text{corr}(r_t^L, r_{t-1}^S)$	−0.036	0.015	0.017	0.054	0.060	0.073	0.078	0.083	0.090	0.111	0.120	0.093
$\text{corr}(r_{t-1}^L, r_t^S)$	0.133	0.060	0.055	0.087	0.085	0.104	0.101	0.101	0.098	0.137	0.139	0.177
$\text{corr}(r_t^L, r_{t-1}^L)$	−0.038	−0.017	−0.017	0.023	0.027	0.039	0.041	0.046	0.048	0.078	0.083	0.068
$\text{corr}(r_t^S, r_{t-1}^S)$	0.140	0.103	0.102	0.137	0.139	0.156	0.158	0.156	0.159	0.189	0.199	0.218

Quotes

	Open	<1100h	>1100h	<1200h	>1200h	<1245h	>1245h	<1330h	>1330h	<1430h	>1430h	Close
$\text{corr}(r_t^L, r_t^S)$	0.901	0.925	0.926	0.919	0.920	0.923	0.923	0.924	0.924	0.919	0.919	0.905
$\text{corr}(r_t^L, r_{t-1}^S)$	−0.041	0.014	0.017	0.054	0.056	0.073	0.077	0.082	0.088	0.110	0.114	0.093
$\text{corr}(r_{t-1}^L, r_t^S)$	0.125	0.058	0.058	0.085	0.084	0.101	0.099	0.098	0.098	0.134	0.137	0.184
$\text{corr}(r_t^L, r_{t-1}^L)$	−0.045	−0.017	−0.015	0.024	0.025	0.039	0.040	0.046	0.048	0.078	0.081	0.074
$\text{corr}(r_t^S, r_{t-1}^S)$	0.125	0.101	0.104	0.133	0.135	0.156	0.156	0.153	0.158	0.189	0.192	0.221

Table 2: **Confidence intervals for autocorrelation patterns of NYSE portfolios sorted by market capitalization** (obtained on annual basis from Datastream). Confidence intervals have nominal 95% coverage and are based on a blocks of blocks bootstrap with a block length of 5 and 9999 bootstrap replications. Only common shares of companies whose primary geographical location is the US are considered. The large (L) and the small (S) market capitalization portfolios each have 250 stocks. Sample period is 1993–1998. $D1$ denotes $[\mathrm{corr}(r^S_{t-1}, r^L_t)_{a\ell} - \mathrm{corr}(r^S_{t-1}, r^L_t)_{b\ell}]$. $D2$ denotes $[\mathrm{corr}(r^L_{t-1}, r^S_t)_{b\ell} - \mathrm{corr}(r^L_{t-1}, r^S_t)_{a\ell}]$. $D3$ denotes $D1 + D2$.

Trades

	1100h		1200h		1245h		1330h		1430h	
	lower	upper	lower	upper	lower	upper	lower	upper	lower	upper
$\mathrm{corr}(r^L_{t-1}, r^S_t)_{b\ell}$	−0.027	0.147	0.016	0.173	0.028	0.172	0.023	0.175	0.066	0.206
$\mathrm{corr}(r^L_{t-1}, r^S_t)_{a\ell}$	−0.038	0.135	0.012	0.168	0.022	0.165	0.017	0.169	0.071	0.211
$D1$	−0.003	0.007	0.002	0.012	−0.001	0.010	0.000	0.012	0.004	0.015
$D2$	0.000	0.011	−0.003	0.007	−0.002	0.008	−0.002	0.009	−0.010	0.004
$D3$	0.001	0.014	0.004	0.015	0.001	0.014	0.003	0.015	0.000	0.014

Quotes

	1100h		1200h		1245h		1330h		1430h	
	lower	upper	lower	upper	lower	upper	lower	upper	lower	upper
$\mathrm{corr}(r^L_{t-1}, r^S_t)_{b\ell}$	−0.030	0.144	0.014	0.170	0.026	0.169	0.021	0.172	0.064	0.204
$\mathrm{corr}(r^L_{t-1}, r^S_t)_{a\ell}$	−0.030	0.144	0.011	0.168	0.021	0.164	0.021	0.173	0.068	0.208
$D1$	0.000	0.006	−0.001	0.006	0.001	0.006	0.003	0.009	0.001	0.008
$D2$	−0.003	0.003	−0.001	0.004	−0.001	0.005	−0.003	0.003	−0.006	0.001
$D3$	0.000	0.006	0.000	0.008	0.003	0.008	0.002	0.009	−0.001	0.005

Table 3: **Autocorrelation patterns of NYSE portfolios sorted by trade frequency.** The frequently-traded portfolio (f) and the infrequently-traded portfolio (i) each have 250 stocks. Only common shares of companies whose primary geographical location is the US are considered. Sample period is 1993–1998. $<\tau$h indicates the last trade prior to τh. $>\tau$h indicates the first trade after τh.

						Trades						
	Open	<1100h	>1100h	<1200h	>1200h	<1245h	>1245h	<1330h	>1330h	<1430h	>1430h	Close
$\mathrm{corr}(r_t^f, r_t^i)$	0.900	0.934	0.938	0.932	0.937	0.930	0.934	0.934	0.935	0.930	0.930	0.916
$\mathrm{corr}(r_t^f, r_{t-1}^i)$	−0.060	−0.021	−0.016	0.023	0.034	0.046	0.052	0.052	0.061	0.085	0.095	0.076
$\mathrm{corr}(r_{t-1}^f, r_t^i)$	0.187	0.111	0.104	0.127	0.121	0.147	0.143	0.141	0.138	0.172	0.175	0.205
$\mathrm{corr}(r_t^f, r_{t-1}^f)$	−0.046	−0.017	−0.017	0.037	0.040	0.057	0.057	0.065	0.065	0.078	0.083	0.080
$\mathrm{corr}(r_t^i, r_{t-1}^i)$	0.154	0.103	0.102	0.119	0.124	0.139	0.144	0.136	0.141	0.189	0.199	0.200

						Quotes						
	Open	<1100h	>1100h	<1200h	>1200h	<1245h	>1245h	<1330h	>1330h	<1430h	>1430h	Close
$\mathrm{corr}(r_t^f, r_t^i)$	0.896	0.937	0.938	0.935	0.936	0.933	0.934	0.936	0.936	0.932	0.932	0.920
$\mathrm{corr}(r_t^f, r_{t-1}^i)$	−0.064	−0.022	−0.015	0.025	0.029	0.047	0.051	0.051	0.061	0.084	0.091	0.080
$\mathrm{corr}(r_{t-1}^f, r_t^i)$	0.176	0.107	0.105	0.123	0.122	0.144	0.142	0.139	0.139	0.169	0.171	0.208
$\mathrm{corr}(r_t^f, r_{t-1}^f)$	−0.054	−0.017	−0.015	0.038	0.038	0.056	0.057	0.064	0.066	0.078	0.081	0.086
$\mathrm{corr}(r_t^i, r_{t-1}^i)$	0.142	0.101	0.104	0.117	0.121	0.141	0.141	0.133	0.141	0.189	0.192	0.204

Table 4: **Confidence intervals for autocorrelation patterns of NYSE portfolios sorted by trade frequency.** Confidence intervals have nominal 95% coverage and are based on a blocks of blocks bootstrap with a block length of 5 and 9999 bootstrap replications. The frequently-traded portfolio (f) and the infrequently-traded portfolio (i) each have 250 stocks. Only common shares of companies whose primary geographical location is the US are considered. Sample period is 1993–1998. $D1$ denotes $[\mathrm{corr}(r^i_{t-1}, r^f_t)_{a\ell} - \mathrm{corr}(r^i_{t-1}, r^f_t)_{b\ell}]$. $D2$ denotes $[\mathrm{corr}(r^f_{t-1}, r^i_t)_{b\ell} - \mathrm{corr}(r^f_{t-1}, r^i_t)_{a\ell}]$. $D3$ denotes $D1 + D2$.

Trades

	1100h		1200h		1245h		1330h		1430h	
	lower	upper	lower	upper	lower	upper	lower	upper	lower	upper
$\mathrm{corr}(r^f_{t-1}, r^i_t)_{b\ell}$	0.025	0.200	0.050	0.207	0.075	0.217	0.066	0.215	0.103	0.240
$\mathrm{corr}(r^f_{t-1}, r^i_t)_{a\ell}$	0.010	0.186	0.041	0.198	0.067	0.209	0.058	0.207	0.109	0.247
$D1$	−0.001	0.011	0.004	0.014	0.000	0.011	0.003	0.014	0.005	0.017
$D2$	0.001	0.013	−0.002	0.011	−0.001	0.010	−0.003	0.011	−0.010	0.003
$D3$	0.005	0.018	0.007	0.019	0.004	0.017	0.006	0.019	0.001	0.013

Quotes

	1100h		1200h		1245h		1330h		1430h	
	lower	upper	lower	upper	lower	upper	lower	upper	lower	upper
$\mathrm{corr}(r^f_{t-1}, r^i_t)_{b\ell}$	0.021	0.197	0.046	0.203	0.072	0.215	0.063	0.212	0.100	0.238
$\mathrm{corr}(r^f_{t-1}, r^i_t)_{a\ell}$	0.017	0.194	0.043	0.200	0.067	0.209	0.063	0.211	0.103	0.241
$D1$	0.003	0.010	0.000	0.007	0.001	0.007	0.006	0.013	0.003	0.010
$D2$	−0.002	0.005	−0.002	0.004	0.000	0.006	−0.003	0.004	−0.005	0.002
$D3$	0.004	0.012	0.001	0.008	0.003	0.010	0.006	0.014	0.002	0.008