Anomalies in the Serial Correlation of Returns

Gautam George *

August 15, 2006

Summary

This is an enquiry in to the structure of equity returns autocorrelation with respect to the distribution of returns. An investigation of autocorrelation conditional on extreme values leads to two stylised facts that reveal anomalous behaviour. First, autocorrelation in the tails of the returns distribution tends to be lower than unconditional returns autocorrelation, and then, the lower tail often exhibits serial correlation of opposite sign: this is statistically inconsistent with standard models of stock returns processes, as shown. Liquidity is identified as an explanatory factor. An absence of liquidity in the market for the asset dampens returns autocorrelation. Liquidity also has a more pronounced effect on serial correlation than volume, hitherto offered as a factor that suppresses autocorrelation. Extreme return events are associated with poor market liquidity, explaining lower tail autocorrelations. A more traditional microstructure view of dealer pricing can also be used to justify lower autocorrelations in less liquid markets.

Is the serial correlation of returns on days of extreme negative return different in any way from that on days of extreme positive return, or on other normal days? Further, if there are singular differences in conditional serial correlation, why might this be so? This preliminary study of the structure of first order returns autocorrelation[1] with respect to the distribution of returns finds that serial correlation in the tails of the distribution is lower than that elsewhere; further autocorrelation in the lower tail is expected to be negative in sign, which is more often than not opposite to the sign of the unconditional autocorrelation of returns.

Yet, this is antithetic to the most commonly employed models of asset price. Long and widely have autoregressive (AR) processes and their variants been used to describe time series of prices. They voraciously pervade the field [4] because they portray the autoregressive features of mean return, as well as persistence in volatility very well. The statistical properties of conditional autocorrelation in AR(1) and GARCH(1,1) with AR(1) mean processes, in particular, serve as a contrary prediction to what we observe empirically, i.e. autocorrelation in the tails should be higher than elsewhere. Thus, the stylised facts presented here

*The author is a doctoral student at Brandeis University and may be reached at ggeorge@brandeis.edu.

[1]Implicitly, all further references to autocorrelation will be to first order autocorrelation in returns, and those to *conditional* autocorrelation will be to autocorrelation conditional on a set of extreme returns.

are anomalous. In explaining this anomaly in serial correlation, market liquidity is identified as factor that potentially retards the extent of serial correlation in two possible ways.

First, the association of large price changes with liquidity motivated trades, rather than with changes in market wide valuations of the stock price or trading motivated by privately held information, may be used to explain this aberration. In general, such large price changes are concomitant not only with larger volumes, but also with poorer market liquidity conditions. Traders on the opposing side of such trades require a discount to buy (or a premium to sell) for the usual reasons of inventory risk and adverse selection. They act as liquidity providers in a market in which there occurs trading activity not related to price adjustment. When they reverse their trades, subsequently, it is therefore at a profit (positive return), causing autocorrelation to be lower than its unconditional value. This is the reason, we venture, for the observation of lower returns autocorrelation in the tails.

Supposing that non-informational trades, motivated by liquidity demands, are signified by high volume, Campbell, Grossman and Wang [9] have already shown such trades to induce a decrease in serial correlation. This owes to the heightened return expectations of trade counterparties who fulfil the role of market maker. They argue that price changes (not necessarily extreme or even large) resulting from a release of information to the public leave expected return unchanged. Also these information related price changes can be distinguished from price changes arising from liquidity trades by the larger volume attendant on the latter. Thus, they search for and identify a negative relation between trading volume and serial correlation of returns. Liquidity, or what they call non-informational, demands for trade in the asset reduce serial correlation. Here, we suggest that it is, instead, the scarcity of available market liquidity[2] represented by the Amihud Illiquidity measure[3] that lowers returns autocorrelation.

We verify an inverse relationship between volume and serial correlation, but detect a much stronger relationship between illiquidity as measured by the Amihud Index and serial correlation. In earlier support of this relationship, Chordia, et. al. [3] have shown, for portfolios sorted by volume turnover and liquidity as measured by the index and held for a week, that high turnover low liquidity portfolios reverse the most. Our data is a broad sample of individual stocks among which numbers one traded index fund. We deliberately avoid using indices to altogether circumvent non-synchronous trading biases in our conclusions. While it has been widely documented elsewhere that volume and absolute return are positively related [22], we find some evidence that so are illiquidity and absolute return; so, on days of extreme return the market is characterised by relatively more adverse liquidity conditions. It follows, then, that autocorrelation will be suppressed on such days.

Any plausible liquidity based justification of these anomalies in autocorrelation must be consistent with traditional microstructure perspectives of the information economics of markets. The uninformed traders or noise traders commonly found in these models induce negative autocorrelation in returns via the Roll bounce mechanism. In contrast, informed trading activity invokes a gradual price adjustment and a resultant positive autocorrelation in returns. In a regret free, risk neutral price setting environment, price changes

[2]In their connotation of the word, liquidity refers to trades not related to new information; in ours it refers to the attribute of the market for the asset.

[3]Henceforth, when market liquidity is referred to, unless the more general sense does apply, it is used in the specific (and implicit) context of the Amihud Illiquidity measure.

occur in multiples of the half spread, and the spread is proportional to the degree of informed activity in the market. The rate of any adjustment of price to true value is therefore dependent on the spread, and the concentration of informed traders. When informed traders abound the spread is larger (on account of larger adverse selection costs) and adjustment is more rapid; consequently returns are less persistent, having lower autocorrelation, than when the concentration of informed traders is lower and the adjustment of price slower. Hence, more liquid markets, viewed as those with less informed activity or a weak concentration of informed traders, might reasonably be expected to be associated with higher serial correlation in returns.

Reversals in extreme price decreases have already been documented in more than one stock market [2, 6, 7, 18, 26]. Periods immediately succeeding those that witness a large negative return show a persistent adjustment in the form of a positive return, which contradicts the traditional concept of the efficient (or complete and immediate) reflection of public information in market prices. In view of unconditional autocorrelations that are weakly positive at the daily frequency, this corresponds to the stylised fact we uncover that left tail autocorrelation tends to oppose in sign unconditional autocorrelation. An accessory contribution of this essay is that in seeking to explain the anomalies in serial correlation structure, we carve a nexus between market liquidity and reversals. It is the scarcity of liquidity, whether on account of heavy information related trading, or unidirectional non-informational trading, that affects autocorrelation in the prescribed manner and hence causes reversals. However, our thrust will remain on enunciating and then explaining serial correlation anomalies rather than a reversal of extreme returns.

There is much work that has concentrated on determinants of the unconditional serial correlation of returns across the entire distribution [12, 14, 24, 25, 28, 30], but our focus here will be the conditional serial correlation in different regions of the returns distribution. This work should be of interest to anyone involved in the valuation of equities and the movement of prices. It is already a commonly held view that falling markets reverse more easily than do rising markets; learning the source of these reversals will surely aid in devising the strategies to potentially exploit what is now verified here as one of two stylised facts of conditional autocorrelation. The role of liquidity trades in lowering serial correlation, and the prevalent relationship with market attributes of illiquidity and high volume might serve in the design of markets. So too might the implications for market overreaction to news, and the adjustment that follows.

The next section proposes two stylised facts of asset returns serial correlation, and shows their aberrance in light of weakly positive daily autocorrelations. The central hypothesis to justify these facts is presented immediately after, along with a brief synopsis of pertinent literature, where necessary. Lastly, we draw up tests, obtain some results and offer some evidence that will support our hypothesis, before we conclude.

1 The Conditional Serial Correlation of Returns

We compare the serial correlation of returns across different return quantile groups, for a range of stocks traded on all three major American stock exchanges, to unearth the two facts below.

(a) Serial correlation in the extreme tails of the returns distribution is *lower* than is the unconditional serial correlation of returns. Thus, if ρ_T represents the serial correlation conditional on contemporaneous return belonging to pooled extreme (upper and lower) quantiles, and ρ stands for unconditional serial correlation, $\rho_T < \rho$.

(b) Serial correlation in the extreme negative tail of the returns distribution is lower than that in the corresponding right tail, or $\rho_{Q_L} < \rho_{Q_R}$, where the two terms represent serial correlation in the lower extreme quantile and upper extreme quantile respectively. Thus, the attribute of lower tail autocorrelation relative to unconditional autocorrelation is more pronounced in the left tail, where the autocorrelation is often of opposite sign so $\rho_{Q_L}.\rho < 0$. If daily returns autocorrelations are presumed weakly positive then conditional autocorrelation in the left tail has a tendency to be negative while that in the right tail is not as often negative.

The data we have used to arrive at these statements is from the Center for Research in Securities Prices. Since this is a well known source of data, we delay a more thorough description of our data set until later in this section. For the immediate moment we merely note that a total of seventy nine returns time series are employed, all of which are for actively traded issues. The stocks in the sample are enumerated in an appendix. We perform two experiments on these series. The first compares sample autocorrelation in pooled tails of up to twenty percent weight with the unconditional sample autocorrelation. The second compares the autocorrelations in left extreme, right extreme and central quantiles of equal weight. First, though the significance of these stylised facts are gauged against two common models of the returns process: the AR(1) model, and the GARCH(1,1) model. They will emerge as contrary to the statistical properties of these standard models.

1.1 Autocorrelation in the Tails of Standard Processes

The incidence of a breakdown in correlation structure in i.i.d. normal variables has already been recorded by Boyer, Gibson and Loretan [5]. They show that for a correlated bivariate i.i.d. series, correlation is higher when conditional on extreme values of one variable than otherwise. Their results are appropriate to the case of multivariate series assumed to follow the strongest version of the random walk with i.i.d. increments. Of course, since the true returns autocorrelation would be zero, we do not consider the random walk model for returns and hence are not in a position to directly apply the results of Boyer, et al., but proceed, instead, in analogous fashion to supplement their theorem for a serially correlated sample.

The first order autocorrelated process is not in the least recalcitrant to analysis: consider the univariate process, $x_t = \alpha + \rho x_{t-1} + \epsilon_t$, $\epsilon_t \sim N(0, \sigma^2)$. Since the data generating process is homogeneous along the entire series, if we were to choose pairs (x_t, x_{t+1}) conditioned on $x_t \in S$, where S is some subset of the state space, possibly but not necessarily the set of extreme tail returns, we would still observe that $x_{t+1}|x_t \in S = \alpha + \rho x_t | x_t \in S + \epsilon_{t+1}$. This does not suggest that autocorrelation conditioned on any event should be the same as the unconditional autocorrelation of the series, however, because stationarity conditions no longer prevail in the conditioned sample space. Referring to the conditional variance of $x_t|x_t \in S$ as σ_c^2, we obtain our first result below as a corollary to Theorem 1 of Boyer et al., though we cannot at all rely on their method of expressing one normal i.i.d. variable as a linear function of the other [5], and so proceed instead from first principle.

Corollary 1 *For an AR(1) series with unconditional first order autocorrelation ρ and errors generated from independent $N(0,\sigma^2)$ distributions, the first order autocorrelation conditioned on realisations $x_t \in S$ is $\rho\Big[\rho^2 + \dfrac{\sigma^2}{\sigma_c^2}\Big]^{-\frac{1}{2}}$, where σ_c is the conditional variance.*

Proof: Corollary 1

$$
\begin{aligned}
Corr(x_t, x_{t+1}|x_t \in S) &= \frac{E(x_t x_{t+1}|x_t \in S) - E(x_t|x_t \in S)E(x_{t+1}|x_t \in S)}{[Var(x_t|x_t \in S)\,Var(x_{t+1}|x_t \in S)]^{\frac{1}{2}}} \\
&= \frac{E(\alpha x_t + \rho x_t^2 + x_t\epsilon_{t+1}|x_t \in S) - E(x_t|x_t \in S)E(\alpha + \rho x_t + \epsilon_{t+1}|x_t \in S)}{[Var(x_t|x_t \in S)\,Var(\alpha + \rho x_t + \epsilon_{t+1}|x_t \in S)]^{\frac{1}{2}}} \\
&= \frac{\rho\sigma_c^2}{[\sigma_c^2(\rho^2\sigma_c^2 + \sigma^2)]^{\frac{1}{2}}} = \rho\Big[\rho^2 + \frac{\sigma^2}{\sigma_c^2}\Big]^{-\frac{1}{2}}
\end{aligned}
$$

A few observations may immediately be made. Only in the trivial case when $\rho = 0$ is the autocorrelation unaffected by conditioning. When the whole (unconditional) sample is considered and consequently, $\sigma_c^2 = \dfrac{\sigma^2}{1-\rho^2}$, then the (unconditional) serial correlation is ρ, as expected for the first order autocorrelated process. Then too, conditional autocorrelation should have the same direction as unconditional autocorrelation does. When the sample is conditioned on pooled tails, $\sigma_c^2 > \dfrac{\sigma^2}{1-\rho^2}$, and hence, the denominator of the expression for conditional autocorrelation is exceeded by unity. Conditional autocorrelation is therefore *higher* in absolute value than unconditional autocorrelation. Allowing that daily returns autocorrelation is positive, conditional autocorrelation should be higher than unconditional autocorelation. To compare the autocorrelations in central and the left extreme quantile, we acknowledge that the conditional variance in the lower tail, σ_l^2, is greater than the conditional variance in the central quantile of equal weight, σ_c^2. This will be true for all well behaved distributions[4] (with median and mode sufficiently close to each other, and near symmetrical shape, in particular). Then, $\dfrac{\rho_l}{\rho_c} = \dfrac{\sigma_l^2(\sigma_c^2\rho^2 + \sigma^2)}{\sigma_c^2(\sigma_l^2\rho^2 + \sigma^2)} > 1$, so that conditional autocorrelation in the left tail is *higher* than that in a central quantile of equal weight. The result applies equally well to the upper tail of an AR(1) process.

The reader may have noticed that the result would also follow from substituting the unconditional variance $\sigma_u^2 = \dfrac{\sigma^2}{1-\rho^2}$ in the result of Boyer et al. for i.i.d. distributions, but this is clearly an invalid route to the same result, in the case of serially correlated sequences.

In a similar manner as for the AR(1) process, it is easy to show that a GARCH(1,1) process will also conform to the general result of our corollary, with one modification. The conditional variance of the innovation replaces its unconditional variance. Specifically, consider such a process which also includes an AR(1)

[4] To constrain to the least extent possible the distribution underlying the returns process, we do not compare the autocorrelation in each tail separately with the unconditional sample correlation: on account of truncated tail observations concentrated around a local mode, the variance in the tail may in some cases prove to be *lower* than the unconditional variance. Such distributions are beyond the realm of assumptions commonly made about the distribution, and so, accordingly, we ignore them.

mean component, $x_t = \alpha + \rho x_{t-1} + \epsilon_t$, where the variance of ϵ_t conditional on the information set Ψ_{t-1} is $\sigma_t = a + b\epsilon_{t-1}^2 + c\sigma_{t-1}$. Then, the conditional autocorrelation is, again, $\rho\Big[\rho^2 + \frac{\sigma^2}{\sigma_c^2}\Big]^{-\frac{1}{2}}$ where σ^2 is the *conditional* variance of ϵ_t and σ_c^2 is the variance of the conditional realisations x_{t-1} as before. It is clear that we expect *higher* tail autocorrelations relative to the unconditional counterpart, and also to that in a more central quantile of equal weight.

To broadly verify these results for the specific condition that x_t lies in the pooled return distribution tails, we simulate an AR(1) series and another GARCH(1,1) with AR(1) mean series, and observe the accuracy with which we predict the pooled tail autocorrelations. It is clear from Figure B.1 in the appendix that our corollary holds to good account. These figures indicate how with falling variance, as the tails increase in size the conditional autocorrelation decreases. It ultimately converges to the unconditional autocorrelation when the pooled tails include the whole sample.

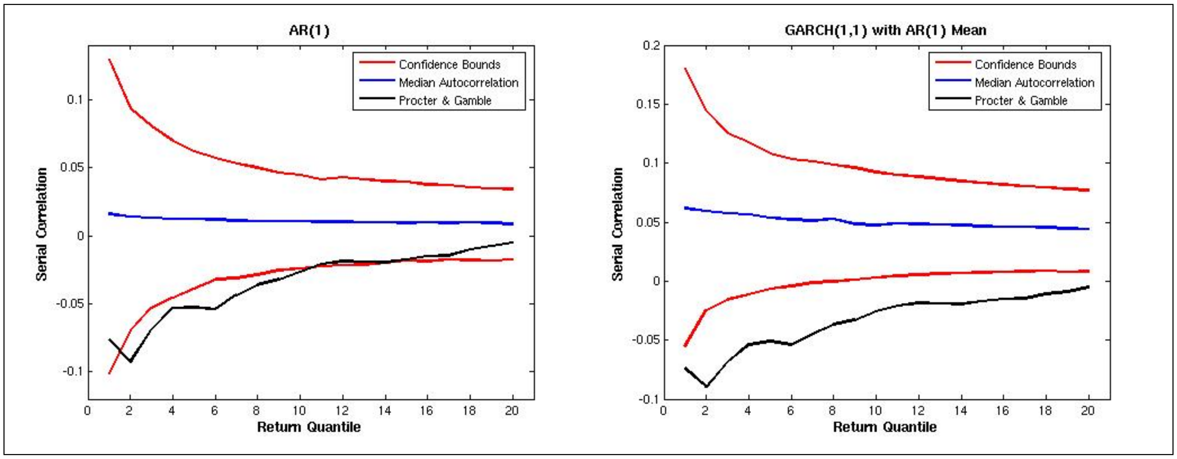
1.2 The Behaviour of Tail Autocorrelation in Stock Returns

Next, we exhibit the evidence gathered in support of our two key stylised facts above. That a statistically significant predilection exists for daily stock price returns to reverse to higher than average levels immediately following a day of excess negative return was first documented by Bremer and Sweeney [7]. Similar reversal phenomena have been observed at higher frequency [18], as well as for less extreme negative return thresholds than used by these authors [8]. Though Bremer and Sweeney conclude that a cumulative average return in excess of two percentage points over the three days subsequent to an extreme price change contradicts weak form market efficiency, other authors have contrarily found that the extent of reversal is actually subsumed by large bid-ask spreads [2].

To emphasise the feature we quest for in this section, we preliminarily compare the autocorrelations in the pooled return tails for one of the stocks in our sample (Procter & Gamble) to their one-sided 95% confidence intervals in Figure 1. The confidence intervals are obtained by one thousand generations of an AR(1) series, and a GARCH(1,1) with AR(1) mean series, each fitted to the actual return time series for the stock being considered. The observed conditional autocorrelations lie below the lower bound over the entire range of tail weights we have employed for the second of these models, and upto the twentieth percentile of pooled extreme returns, for the first. This encourages a more thorough estimation of the conditional autocorrelations for the whole sample of seventy nine stocks.

So, turning to the stock return series now, we undertake an inspection of autocorrelation in the lower and upper tails. We use a sample of seventy nine traded assets, seventy eight of which are common stock and one which is a traded index fund. The seventy nine time series are panelled into three groups: the first consists of the thirty stocks of the Dow Jones Industrial Average (DJIA) Index and the S&P 500 traded index fund; the second consists of the thirty two stocks of the S&P 500 Index with a market capitalisation of over $50 billion not included in the first group; and the last includes all individual stocks that were studied by Campbell, Grossman and Wang [9] and not included in the first two groups. These stocks are chosen because they are among the most actively traded issues on stock exchanges, with very low spreads to minimise the bid-ask

Figure 1: Pooled Tail Autocorrelations for Procter & Gamble



bounce effects that commonly adulterate inferences about returns autocorrelation. The data ranges over the period from July 1962 to December 2004 for the longest series, but many series are not as long, leaving unbalanced panels. As mentioned before the source of data is the Center for Research in Securities Prices and our principal fields of interest are total number of shares outstanding, daily volume, close prices and returns (based on close prices).

Each daily return series is first sorted by contemporaneous return into a list of pairs (x_t, x_{t+1}), and then autocorrelations in the relevant quantiles estimated. The Fisher transformation is used to obtain bounds on the correlation estimates. We compare pooled tail with unconditional sample autocorrelations and also compare estimates of correlation in each tail to one another. Second, we study autocorrelation in the extreme quantiles and the central quantile. Our findings are tabulated in Tables 1 and 2. The first two columns of the first table report the incidence (as a proportion of the sample size) of the two anomalies in serial correlation: a reversal in the sign of conditional autocorrelation in the tail ($\rho_T \rho < 0$), and a tendency for it to be below the unconditional autocorrelation ($\rho_T < \rho$). For example, the first entry in Table 1 indicates that of the seventy nine stocks surveyed, 48.10% demonstrate a sign change in conditional autocorrelation in pooled tails of 2% weight each. The probability that the incidence rate is lower in a sample of seventy nine return series all following the GARCH(1,1) form with an AR(1) conditional mean is given in parenthesis for each entry. These probabilities have been constructed from one thousand simulations[5] of the assumed process for returns. The complement of each informs us of the probability that as high an incidence rate may occur by chance in a sample size as large. Thus, the first entry in the table also tells us that it is virtually impossible that 48.10% of the sample may exhibit a sign change in conditional autocorrelation in pooled tails of 2% weight each completely by chance. The next two columns consider the sign changes in unpooled tail autocorrelations up to 20% tail weight: left tails exhibit autocorrelation sign changes relative to unconditional autocorrelation more obviously than right tails. On the contrary at higher weights the right tails show a tendency to have the *same* sign as unconditional correlation; this is evinced by increasing levels of significance for the incidence rate of sign changes which approach zero as the tail width increases. The

[5]The one thousand simulated series are generated using coefficients fitted to the Procter & Gamble series.

last column shows the clear tendency of the lower tail autocorrelations to be exceeded by the upper tail autocorrelations ($\rho_L < \rho_U$).

Table 2 contains the results of the second experiment. Lower tail, as well as upper tail autocorrelations are more frequently *lower*, not higher, than the autocorrelation of the median-centred quantile of equal weight ($\rho_T < \rho$). This effect is more pronounced for the left tail, than the right. Also, sign changes in autocorrelation are highly significant in the lower tail, whereas the upper tail experiences *no* changes in autocorrelation sign to high significance.

Thus, it stands out clearly that rather than exceeding unconditional autocorrelation or conditional autocorrelation towards the median, the magnitude of autocorrelation conditional on extreme returns actually tends to be superseded by them (so that $\rho_T < \rho$), and also this tendency reduces as the tails gain cumulative density. Further, a significantly large number of lower tail autocorrelations reverse sign relative to unconditional autocorrelation, whereas the same is not as strikingly true of upper tail autocorrelations, which, consistent with this feature show a clear tendency to exceed the former[6]. It is also evident that there are an inordinately large number of sign changes in lower tail autocorrelation with respect to central quantile autocorrelation. Not only is this contrary to the admitted characteristic of both the standard processes shown earlier, but also the simulations rebut almost all doubt that this may in fact be a mere product of chance. Based on our estimates of autocorrelation, we evaluate the probability that pooled tail autocorrelation is lower than unconditional autocorrelation, and the probability that lower tail autocorrelation is exceeded by upper tail autocorrelation, for each return series in the sample. The appendix contains Figures B.2 and B.3, which depict that most stocks in the sample have lower autocorrelations in pooled tails rather than unconditionally with probability 0.80 or more, and also that lower tail autocorrelation is more likely than not to be exceeded by upper tail autocorrelation for a majority of stocks in the sample.

Table 1: Pooled Tails and Comparison of Return Tail Serial Correlations

	Pooled Tails		**Unpooled Tails**		
Quantile	$\rho_T.\rho < 0$	$\rho_T < \rho$	$\rho_L.\rho < 0$	$\rho_U.\rho < 0$	$\rho_L < \rho_U$
2.00%	0.4810 (1.0000)	0.8354 (1.0000)	0.6076 (0.9999)	0.4810 (0.7517)	0.6329 (0.9966)
5.00%	0.3418 (1.0000)	0.8608 (1.0000)	0.6329 (1.0000)	0.4304 (0.6302)	0.7722 (1.0000)
7.50%	0.2405 (1.0000)	0.7722 (1.0000)	0.6582 (1.0000)	0.3797 (0.4614)	0.7722 (1.0000)
10.00%	0.1772 (1.0000)	0.6835 (1.0000)	0.6962 (1.0000)	0.3797 (0.7355)	0.8228 (1.0000)
15.00%	0.1013 (1.0000)	0.6329 (1.0000)	0.7089 (1.0000)	0.3797 (0.8549)	0.8228 (1.0000)
20.00%	0.0506 (1.0000)	0.4810 (1.0000)	0.6835 (1.0000)	0.3418 (0.8245)	0.8608 (1.0000)

The figures in parentheses give the probabilities of observing a lower incidence rate in an AR(1) GARCH(1,1) process, than is actually observed in the sample of seventy nine series. All seventy nine stocks have a pooled tail variance that is higher than the unconditional variance at each quantile size.

Daily returns autocorrelations are thought to be weakly positive, and since conditional autocorrelation should be of the same sign as unconditional correlation, we expect autocorrelation conditioned on the tails to be

[6]This corroborates an earlier finding [2] that stocks reverse to a weaker degree after a day of extreme positive return than after a day of extreme negative return.

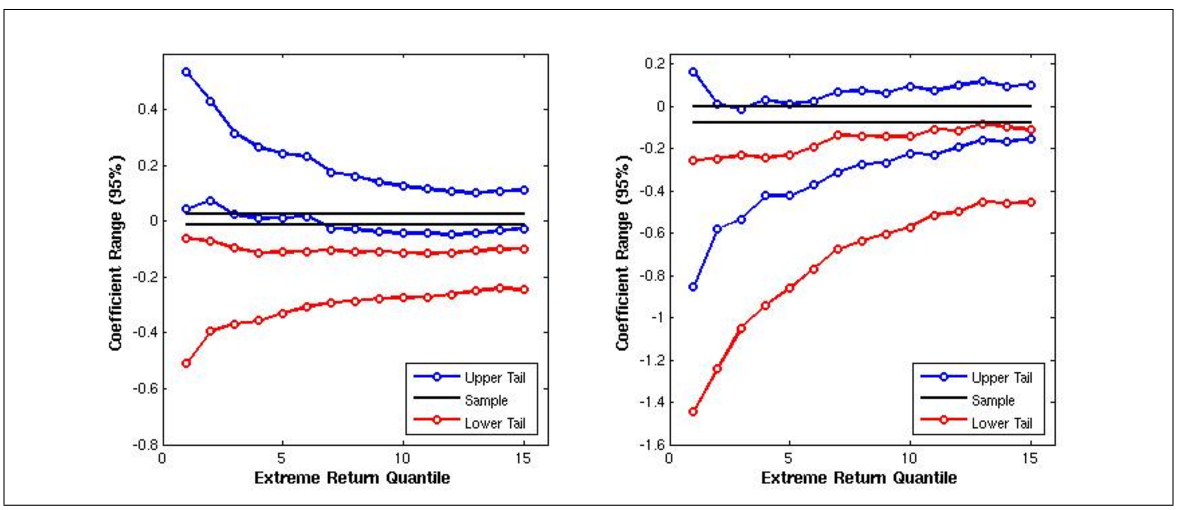
Table 2: Autocorrelation in the Extreme Return Tail and Central Quantiles

	Lower Tail		Upper Tail	
Quantile	$\rho_T.\rho < 0$	$\rho_T < \rho$	$\rho_T.\rho < 0$	$\rho_T < \rho$
10%	0.5775 (1.0000)	0.8873 (1.0000)	0.2958 (1.0000)	0.6479 (1.0000)
15%	0.6479 (1.0000)	0.8873 (1.0000)	0.3803 (1.0000)	0.6056 (1.0000)
20%	0.6479 (1.0000)	0.9014 (1.0000)	0.3521 (1.0000)	0.5493 (1.0000)
25%	0.7324 (1.0000)	0.9296 (1.0000)	0.3239 (1.0000)	0.6620 (1.0000)

The figures in parentheses give the probabilities of observing a lower incidence rate in an AR(1) GARCH(1,1) process, than is actually observed in the sample of seventy one stocks. Eight stocks are dropped because their smallest central quantiles show too little variation to edify. All seventy one of the other stocks have lower central quantile variance than (upper or lower) tail variance. Also, the tail weights used in this table do not assume more extreme values (such as 2%) because of highly restricted variation in the central quantiles of equal weight.

positive for the most part as well. Yet, it has emerged that this is not the case at all; moreover, conditional autoregressions at varying lower tail weight for the stocks that violate this rule show that the negative conditional autocorrelation actually drifts towards the positive quadrant as the tail weight is increased. These regressions do not include weekday or other effects, but are of the simple form $r_{t+1} = \alpha + \rho r_t$, as the attempt is to emphasise the coefficient of correlation more than anything else. The bounds for coefficients of the autoregression in the lower and upper tails for Procter and Gamble and the traded index fund illustrate this point in Figure 2. Attendant here is also is the absence of this attribute for the upper tail coefficients, though it is also obvious from the much wider widths of the confidence bands for upper and lower tail correlation that the conditional sample suffers from inadequate size.

Figure 2: Autoregression in the Return Tails for Procter and Gamble and the S&P 500 traded index fund



This section has demonstrated a tendency of tail autocorrelations to be lower than the unconditional autocorrelation, or that in a central quantile. This phenomenon is clearer in the lower tail which also frequently reverses the sign of unconditional autocorrelation. These empirical regularities are contrary to what widely applied models of stock return suggest for the relative magnitudes of tail autocorrelations.

2 Liquidity and Autocorrelation

This section will introduce the hypothesis that market liquidity is a factor responsible for the peculiar behaviour of autocorrelation in the return tails. First, we extend the antecedent offered by Campbell, Grossman and Wang to build a plausible case. Next, we delve into the microstructure of the market to understand the mechanism by which liquidity may have the proposed effect. The concept of liquidity is rooted deeply in microstructure and closely associated with the information games played by market participants, and so this is a crucial facet of the discussion.

2.1 Non-Informational Trading and Liquidity

In seminal work, Campbell, Grossman and Wang have shown volume, as indicative of non-informational trades, to lower autocorrelation [9]. They use daily price returns to a value weighted index of NYSE and American Stock Exchange stocks over a period of twenty eight years as their primary data set, and a weekday effect on autocorrelation to directly relate higher trading volume to lower autocorrelation. Second order autocorrelations show a similar, if slightly weaker and more complicated, relationship. At higher order autocorrelations even if the effect seems to taper, it certainly does not offset the result. Different measures of volume withstand the same tests of this negative relationship, and so does a sample of thirty two individual stocks, used to escape interference from non-synchronous trading. Their explanation, as mentioned earlier, is that non-informational trades are countered by parties who demand a fee for the liquidity services they provide (as market makers, in a sense [17]), which they charge by means of an adverse price for the liquidity trades. Thus, there is a cost of transacting high volume in a market where a unilateral change of beliefs does not justify this. Expected returns change with a price fall in case of a non-informational sale or a price rise in case of a non-informational purchase, to allow liquidity providers a fee. This price change generates a return in the current period negatively correlated with the expected return realised in the next period. We may view the decrease in subsequent autocorrelation as the obvious fallout of the realisation of this transaction cost.

In a model of rational investors facing uncertainty arising from future dividends and from a noisy signal pertaining to these future dividends, these authors further press their case. They allow investors' risk aversion to vary with time and use this as the reason behind non-informational trades. Stock price can fall either due to a signal describing future cash flows or due to a change in risk aversion levels of a subset of investors. If there is portent of lower future cash flows, this is market wide since the signal is common to all investors, and there is no abnormal trading activity. On the other hand, if some investors develop a greater aversion to risk than they already have, they decide to dispose of their holdings of the asset and this generates increased trading activity. To induce other investors to expand their portfolios in favour of the risky asset, expected return has to rise and this is facilitated by a fall in stock price leading to a negative current return. Of course, since the signal has not indicated lower future earnings, during the next period the stock price returns to its true value (unless the next period involves liquidity trades once again). Therefore, autocorrelation is depressed by these non-informational trades.

The effect of trading volume on returns serial correlation has subsequently been tested by Conrad, Hameed and Niden [12], who find that trading activity does play a significant role in determining returns. Using a sample of NASDAQ stocks they find that return autocovariances are negative in the wake of heavy trading activity, but when such activity declines, these autocovariances turn positive.

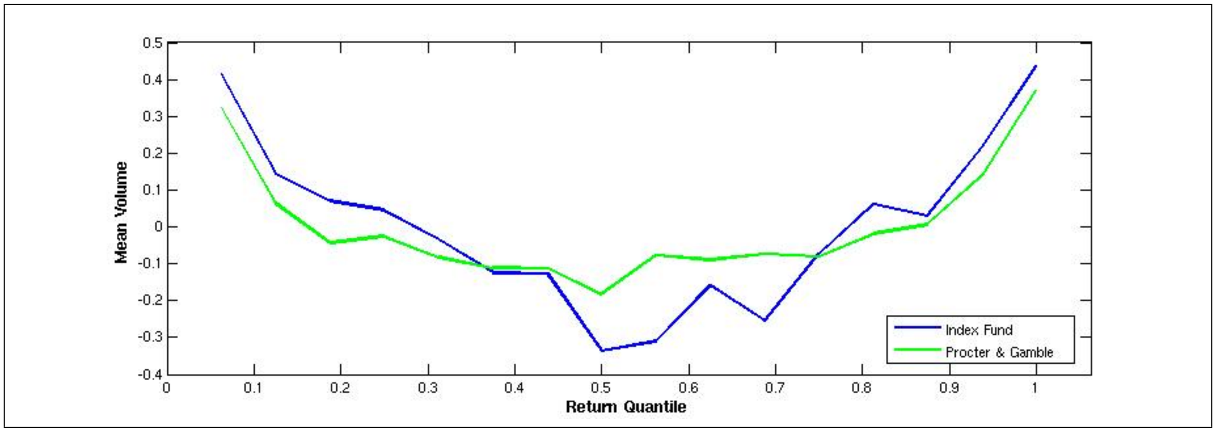
The question naturally materialises, next, whether the stylised facts of extreme returns autocorrelation can be explained by this liquidity demand effect or not[7]. We will also investigate whether existing market liquidity conditions may justify these facts. Somewhat along these lines, Cox and Peterson [13] have earlier tested whether the reversal effect is more pronounced in exchanges such as the American Stock Exchange, where less liquid stocks trade, in comparison to other exchanges; they do not find support for this claim. Though they do demonstrate that smaller (less liquid) stocks reverse more than do larger stocks, this changes once the bid-ask spread is compensated for. Their conclusion is that size more accurately stands in for the spread than for other general liquidity attributes of the market. The reversal experienced on the first day has been explained as a microstructure effect of the Roll bid-ask bounce [29]. However as Cox and Peterson have noted, effects stemming from the bid-ask bounce are not distinct from liquidity effects, as the bid-ask spread itself is a liquidity indicator. Immediately lacking a solution to the possible bounce error, we thwart it to the best of our ability by choosing only large capitalisation stocks which have relatively high liquidity, with spreads of a single tick, in normal market conditions[8].

The connection between volume and returns is relevant to the question of how the liquidity demand effect of Campbell, et al. is related to subjugated serial correlation in return distribution tails. There has been much work on the relation between market returns and trading volume that finds absolute returns to be positively related to volume [11, 15, 27, 32], and there is also abundant evidence that this relationship holds for equities at the daily frequency [19, 20, 31, 33]. This hints that a larger component of the trades that generate extreme returns are non-informational in nature, and cause lower tail autocorrelations, generally. To verify this, we attempt to relate volume turnover, detrended by linear fit, to the different return quantiles. Suspicions that trades resulting in extreme price returns are associated with higher levels of volume on average are confirmed by Figure 3, which plots the average volume turnover level at each return quantile for a typical stock and for the traded index fund. This is not at all surprising in the foreground of the widely accepted relationship between volume and return mentioned. Also, earlier work showing that the relationship between the square of returns or its variance and volume is positive [10, 27] is in harmony with its non-linear shape.

[7] It has been hazarded that the reversal phenomenon is related to block trades and, therefore, their influence on market liquidity [7] but this has not actually been proven.

[8] We offer a heuristic metric of the severity of this problem in our data. For tail returns (upto 20% weight) we use the close bid and ask prices for ten of our stocks to predict reversal, and note that in over 90% of the cases the bounce does not predict more than half of the reversal, as a proportion of it; then we regress the observed reversal on a constant and this predicted reversal. We find that the estimated coefficient of the predicted value is almost never significant, suggesting, therefore, that the bounce has ambiguous effect. The ten stocks we choose here are all NASDAQ stocks as complete closing bid and ask data is available for stocks trading on this exchange in the CRSP database. Generally, spreads on the NYSE are even lower, and so we conclude that for our data least, this problem is not a significant concern.

Figure 3: Volume and Returns



To answer the question at hand then, the volume effect of Campbell Grossman Wang is likely responsible for the anomaly in conditional autocorrelation: extreme daily price changes are frequently caused by non-informational trades, as identified by their association with larger daily volume turnover, and so autocorrelation in the tails is lower. Therefore, autocorrelation in the tail is depressed. In tandem with the positive relationship between volume and absolute return, it would seem that non-informational trades are mostly associated with large price changes and not with more moderate ones. This is not necessary once the state of market liquidity for the asset is substituted for volume in the picture thus far, as we now explain.

As a more direct explanation for anomalous conditional autocorrelations, on the supposition that episodes of extreme price return are characterised by relatively poor market liquidity conditions, next, we seek a relationship between market liquidity and autocorrelation. To encapsulate the aspect of liquidity, we use a variant of the Amihud Illiquidity Index [1]. More details on this measure may be found in the comparative treatment of Hasbrouck [21]. Its selection here is based on its prior use in the literature [3] and also on easy access to volume and return data[9]. This index is defined here as the price impact of each dollar unit of the stock traded over a period, as in Equation 1 below, where r_t, p_t and V_t denote the return, closing price and trading volume on day t respectively.

$$I_t = 10^2 \frac{|r_t|}{\log(p_t V_t)} \tag{1}$$

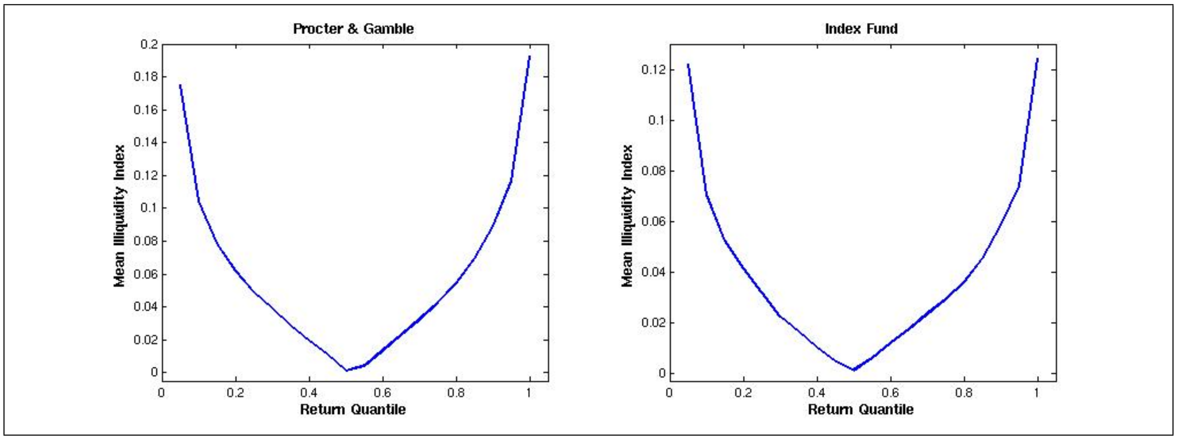
Using this illiquidity measure, Avramov, Chordia and Goyal [3], demonstrate that portfolios of high volume, low liquidity stocks exhibit both larger magnitudes of extreme price changes as well as more negative autocorrelations, so that reversal is more pronounced in them. They treat this as support for the model of non-informational trades. Whereas they use the illiquidity index to classify stocks into portfolios, we use it to capture variation in the daily liquidity conditions for a *particular* stock. This has the advantage of controlling the effects of any idiosyncratic factors besides permitting us to associate the state of temporal liquidity for each *specific* asset to returns autocorrelation.

Alongside our verification of the prevalence of this relationship within the data set under exploration, we also portray in Figure 4 the relationship between the liquidity measure and each returns quantile. This

[9]For effective spread estimates, daily data is not accurate, nor available in the database in use.

figure indicates that our supposition of a positive relationship between liquidity and absolute return is justified. Note that, quite unlike the definition of the illiquidity measure would seem to have it, volume and illiquidity are positively correlated. This may be interpreted thus: when trading volume is relatively low, liquidity for the asset is high because the market has an underutilised absorption capacity for the asset, whereas when trading volume is relatively high the absorption capacity is more overextended than usual, and liquidity is low. An inverse relationship between market liquidity and autocorrelation is consistent with non-informational trades that *do not* result in large price changes: such large volume trades occur on days when the market is highly liquid, so that the return impact is not large.

Figure 4: Liquidity and Returns



Our central premise, that we shall test, is that it is market illiquidity which has a negative effect on the serial correlation of returns, rather than volume, whose measured negative influence has been as that of a proxy variable for liquidity demand. The direct relationship between absolute return and illiquidity would then imply lower tail autocorrelations. Further, we hypothesise that either liquidity itself acts more strongly as a deterrent to autocorrelation in the left tail than in the right tail, or there is asymmetry in the relationship between liquidity and return. Since illiquidity would depress serial correlation to varying extents in the tails, either surmise can explain the tendency for lower tail autocorrelation to be more negative than upper tail autocorrelation.

As a pilot test, we modify the specification of Campbell, Grossman and Wang to include a liquidity term[10]. Their regression form explained daily return with the help of a weekday varying lagged return effect and interactions between volume and return, each lagged by one period. We compare results from a straightforward test of their specification to results from similar regressions that include a liquidity interaction term with and without a volume interaction term. We organise their sample of thirty two stocks into four groups by date, to preserve the faithfulness of the replication here. This exercise allows us an idea of the relative performance of the liquidity measure as an explanatory variable in the original context. The various forms are described in Equations 2, 3 and 4 below.

[10] To be precise, this is the illiquidity measure, rather, but we refer to it as liquidity, equivalently, since there is no risk of ambiguity arising.

$$r_{t+1} = \alpha_0 + (\sum_{i=1}^{5} \alpha_{1,i} D_{i,t} + \beta V_t) r_t \quad (2)$$

$$r_{t+1} = \alpha_0 + (\sum_{i=1}^{5} \alpha_{1,i} D_{i,t} + \gamma L_t) r_t \quad (3)$$

$$r_{t+1} = \alpha_0 + (\sum_{i=1}^{5} \alpha_{1,i} D_{i,t} + \beta V_t + \gamma L_t) r_t \quad (4)$$

Our assay of the preliminary success of the market liquidity effect regressions is based on the relative performance of the volume and liquidity variables. To this end we compare the actual number of stocks that exhibit significant coefficients for liquidity and volume in all three hundred and eighty four (3 x 32 x 4) regressions. The regressions employ White's correction for heteroscedasticity, and the results are exhibited in Table 3 below. It reports the number of regressions (or stocks) in each set of thirty two, for each of the four sample subsets, that possess negative and significant liquidity and volume coefficients. It is clear that liquidity has a wider impact on autocorrelation than does volume, in the individual as well as the joint regressions. Not only are a larger number of coefficients of the expected negative sign for liquidity than for volume, but also, a larger number of coefficients are significantly negative.

Table 3: Pilot Tests of the Liquidity Effect

Sample	*Regression*	$t_\beta < -1.64$	$\beta < 0$	$t_\gamma < -1.64$	$\gamma < 0$
	2	19 (12)	29 (31)		
A (7/62-9/87)	3			25	30
	4	8	23	22	30
	2	11 (11)	26 (30)		
B (7/62-12/74)	3			16	28
	4	3	18	14	28
	2	7 (6)	26 (25)		
C (1/75-9/87)	3			20	31
	4	1	12	19	31
	2	10 (5)	29 (31)		
D (7/62-12/88)	3			23	31
	4	1	11	21	31

The figures in parentheses provide the number of stocks out of the sample size of thirty two that were originally reported [9] as belonging to the specific column they appear in. Slight differences in treatment of the data might result in the small discrepancies recorded.

We consider this sufficient preliminary evidence that returns autocorrelation is negatively influenced by the state of liquidity of the market. In the next section, we introduce our more extensive data set, and elaborate further tests to establish this relationship.

2.2 Informed Traders and Liquidity

The role of information is key in deliberating the systematic influence liquidity in the market for an asset has on the serial dependence of its price returns, since it is against information that the dealer must protect his function of facilitating market liquidity by determining prices at which buy and sell transactions occur. The presence or absence of information as well as its degree of symmetry in the market are distinct in their effect on returns autocorrelation. What follows draws on classic microstructure concepts rooted in Bagehot's observation that dealers (or suppliers of liquidity, in general) levy a spread to compensate not only for the cost of transaction but also for adverse selection in transaction with possibly better informed entities. The widespread concept of liquidity as Kyle's *depth* corresponds directly to the Amihud measure of liquidity [23]. According to this, liquidity is proportional to the trading volume required to impact price by one unit, and is high in markets with a large number of noise traders. These are relatively uninformed traders participating in markets along with informed traders. When markets have a high concentration of informed traders they are comparatively illiquid. Though noise traders in this context share the attribute of being uninformed with the non-informational traders of Campbell et al. there is one important point of departure. Their concerted activity always results in a net order flow of zero since they trade in an uncorrelated manner; if such were the case with non-informational traders, then demand for liquidity would be on opposite sides of the market. Thus, the supply of liquidity to sellers demanding it would issue from the (roughly equal) numbers of buyers also demanding liquidity. All the demand for liquidity would be met from *within* the group of non-informational traders, eliminating the premium or discount. Non-informational trading is correlated and unidirectional, while noise trading is uncorrelated and results in zero net order flow. We rely on the Glosten and Milgrom model of adverse selection [16] to make inferences about returns serial correlation based on the state of information available in the market.

2.2.1 Persistence of Price on Informed Activity

In their model, there are two breeds of the trader that dealers entertain: there are uninformed traders who have access to information *after* its public release and there are informed traders (Kyle's insiders) who possess information *before* release to the public. Though dealers do not know which category the next trader belongs to, from the observation of past prices they are aware of: (a) all public information available in the market, and (b) the relative concentration of informed traders in the market. They use this information to determine what the price markup should be to offset the risk that the next trader is informed. Since all trades with informed traders result in a loss to the dealer, the component of spread that owes to adverse selection risk is lower in markets with a greater proportion of the informed among traders. In such markets the dealer observing lower informed trader activity adjusts transaction and inventory costs by a lower amount for adverse selection in setting spreads than in other markets.

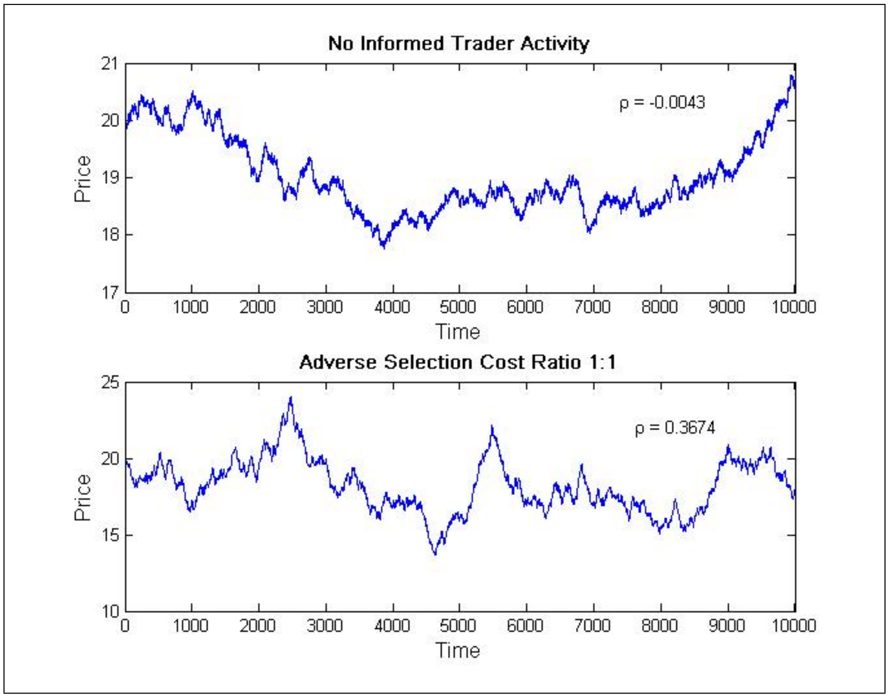
Now, consider the occurrence of an event, knowledge of which is at first available only to the set of informed traders *before* release to the markets. Gradually, more informed traders transact with the dealer, and as price locks into a slow trajectory to its expected value inclusive of the information of the event, returns autocorrelation will tend to be higher than at other times. Ultimately, when the event is announced, or enough informed traders have implicitly given news of it away, this incremental price adjustment ceases. Thus, the role of informed traders is to induce persistence in returns. On the other hand, the activity of uninformed (or non-informational) traders is to induce negative autocorrelation. The relative concentration of informed traders in the market is of direct significance to the magnitude of the positive effect informed trading has on returns autocorrelation. As explained, the larger the adverse selection component of spread, the higher is the dealer's estimate of informed trader concentration in the market. When informed traders comprise a large segment of the market, spreads are higher, and the more quickly price will converge to expected value conditional on all public *and* private information, which we refer to as true value. To understand this, remember that the dealer sets the spread so that he charges no-regret prices for trades. This means that the dealer adjusts his expectation of asset value to the ask price in the event of a buy trade, or the bid price in the event of a sell trade. Thus, when spread is higher the speed of convergence to true value is more rapid. On the other hand when spreads are narrower, price drifts more slowly to true value. Since this period of informed activity is what induces relatively higher autocorrelation as price persists towards the true value, the rate at which it does so is instrumental in the degree of induced positive autocorrelation. As the concentration of informed traders increases, therefore, the positive effect on autocorrelation *reduces*. While the effect of informed activity is to raise autocorrelation, the extent to which it does so, decreases with strengthening informed trader concentration. To summarise, autocorrelation has a tendency to be more negative (or less positive) when trading is non-informational in markets with few information asymmetries.

2.2.2 The Effect of Informed Trading on Returns Autocorrelation in Relief

In this section we exhibit the microstructure effect on autocorrelation of informed trading in a market made by a risk neutral agent cognisant of two categories of market making costs. These costs are: (a) adverse selection costs, or the cost incurred by assuming the role of a relatively *less* informed party to a transaction, and (b) transaction costs including inventory carriage and commissions. Therefore, the spread comprises the adverse selection component of Glosten and Milgrom and also a second component invariant to the activity of informed traders. The concentration of informed traders is known to the market maker, and weighs the probability that the next transaction is information motivated: the higher the concentration of traders the higher is the adverse selection component of spread set by the market maker. Information events are assumed to occur independently of the concentration of informed traders in the market. Admissibly, information events may occur more frequently when there are a greater number of informed traders to root these out, but this is a simplification we adopt here to emphasise the autocorrelation effect described. This may be reconciled with reality by considering only relatively significant events (or those that cause relatively appreciable price changes), which should be less related to the concentration of informed traders than less significant events. Periodically, information becomes available and the informed traders submit orders until the price adjusts to its information-adjusted value. This is exogenously set to a fixed multiple of the current price. A greater number of the trades that take place owe to noise trader activity with a purchase being as likely as a sale. The market maker sets spreads so that the transaction price is regret free, implying that the absolute price change from one to the next trade is always by the exact distance of the half-spread.

When there is a high concentration of informed traders the spread is wider, on account of the variable adverse

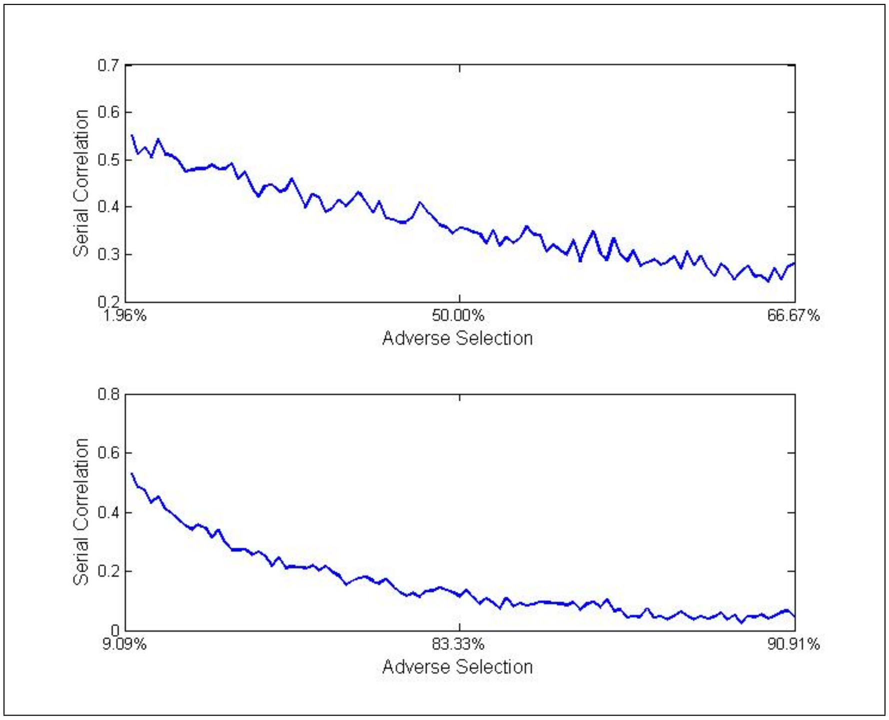
Figure 5: Price Series in Two Markets With Different Informed Trader Concentrations



selection component, and convergence to true value is quicker (in transaction time) in contrast to when there are a smaller concentration of informed traders. The induced spell of higher autocorrelation is longer lasting when there are few informed traders, and the net effect is that returns autocorrelation is higher in the presence of fewer informed traders. Though elementary, this market system depicts the effect of information on autocorrelation, as computational results verify. Figure 5 plots transaction price in two markets: one has no informed traders (so that spread consists of transaction costs alone and no adverse selection costs), the other has exactly half the spread contributed by adverse selection costs. It may be observed that there is more oscillation in price in the first market, leading to the expected negative serial correlation that noise trading causes. In the second market there are clearer price runs and a higher returns autocorrelation. The next figure, Figure 6 shows how autocorrelation varies with the proportion of spread comprised by adverse selection costs. It is clear from either of the two ranges for which plots are exhibited that as the adverse selection component of spread rises, or the concentration of informed traders in the market rises, the serial correlation of returns falls.

Recognising that liquidity is directly borne upon by the information that traders bring to the market place, it is necessary to make a case from this standpoint as well. Viewing illiquidity as a function of the degree to which informed traders occupy the market, the claim that autocorrelation is impeded by poor market liquidity is tenable.

Figure 6: Autocorrelation and Spread at Different Informed Trader Concentrations



3 Tests and Results

The data set has already been described in an earlier section; the exploratory tests of conditional serial correlation employed the same data. Yet, it will only now be appreciated why the thirty two stocks from the individual stock tests of Campbell, Grossman and Wang are included. In comparing the success of illiquidity to that of volume as a determinant of serial correlation, these data samples provide reference to their results.

We do not construct and use indices of any sort so as to avoid the trap of non-synchronous trading. Trading of only a subset of the stocks in the index on a particular day induces measured autocorrelation in returns to be higher than its true value. On another day when a larger subset of the index stocks trade, the autocorrelation would more closely reflect its true value, and would therefore be lower than on days when fewer stocks trade. Yet traded volume on days when large subsets of index stocks trade is higher than that on says when smaller subsets trade, leading to the conclusion, in fact yet unsubstantiated, that volume has a negative influence on autocorrelation.

Descriptive statistics for the time series are available in the appendix. The explanatory variables, volume turnover and liquidity take zero and negative values, because we have not filtered out zero volume observations from the data sample. Days on which no trading occurs and yet price moves are recorded are possible[11], since changes in market valuation occur without concurrent trading activity. Such days, the volume is taken to be a small number of the order of 10^{-3} so that the illiquidity measure takes negative (and not undefined)

[11] They represent an insignificant proportion of the data, however.

values, and volume turnover is zero.

The specification of tests in this section departs from the pilot in two respects: we include a weekday constant effect rather instead of a weekday return effect, as this is conventional; second, and more importantly, we separate the effects of volume and liquidity in each of the two relevant return quadrants, to judge whether these factors have disparate effects conditional on return sign. If so, the second stylised fact of more negative autocorrelation in the left tail in contrast to the right tail might be explained by the stronger negative influence of illiquidity in the left tail. Errors are assumed to be heteroscedastic though uncorrelated, so that White's correction is employed as before. The complete specification is provided in Equations 5, 6, and 7 below.

$$r_{t+1} = \sum_{i=1}^{5} \alpha_{0,i} D_{i,t} + \Big(\alpha_1 + \beta_1 V_t \mathrm{I}(r_t < 0) + \beta_2 V_t \mathrm{I}(r_t > 0)\Big) r_t \tag{5}$$

$$r_{t+1} = \sum_{i=1}^{5} \alpha_{0,i} D_{i,t} + \Big(\alpha_1 + \gamma_1 L_t \mathrm{I}(r_t < 0) + \gamma_2 L_t \mathrm{I}(r_t > 0)\Big) r_t \tag{6}$$

$$r_{t+1} = \sum_{i=1}^{5} \beta_{0,i} D_{i,t} + \Big(\alpha_1 + \beta_1 V_t \mathrm{I}(r_t < 0) + \beta_2 V_t \mathrm{I}(r_t > 0) + \gamma_1 L_t \mathrm{I}(r_t < 0) + \gamma_2 L_t \mathrm{I}(r_t > 0)\Big) r_t \tag{7}$$

A dummy, D_i, is included for each weekday. Coefficients for volume turnover conditional on negative and positive returns are denoted by β_1 and β_2 respectively, while those for illiquidity are denoted by γ_1 and γ_2 respectively. As with the pilot, we expect to find that the coefficients of liquidity as well as volume are negative, with those of liquidity being so more often of the two. Moreover, the coefficients associated with the negative returns should be *lower* than those associated with the positive quadrant. This would imply that each variable acts to the greater detriment of autocorrelation in the negative returns quadrant, as an explanation for lower left tail autocorrelations. The results of a typical regression are exhibited in Table 4, where the coefficients are reported in each column along with the t-statistic for that coefficient in parenthesis, against the relevant regression specification. It is plain that the estimated coefficients for liquidity turn out exactly as expected. One of the individual regression volume coefficients is actually positive, while both the volume coefficients are positive in the joint regression. Also, liquidity appears to be a stronger detriment on autocorrelation in the negative quadrant than in the positive, suggesting an asymmetrical effect. What may also be noticed is that the coefficient of determination (which is abysmal as typical with regressions of this nature) actually rises when liquidity is the explanatory variable. On the whole, the Wald criteria suggest that the regressions become more significant when the illiquidity measure is included.

Next, results for the entire sample are presented in Table 5. This table reports the incidence rates for negative and significant coefficients of interest in each regression. Each column has an entry for the proportion of negative coefficients, and another entry in parenthesis for the proportion of negative and significant coefficients, in that sample group for that regression. For example, the first entry shows that the volume coefficient for negative returns (β_1) in the first regression (Equation 5) is negative for 87.10% of the total number of series in the first sample group of thirty one stocks; the next (in parenthesis) shows that the same coefficient is significantly negative for 22.58% of the total number of series in the same sample group. This

Table 4: Regression Results for IBM

Regression	5		6		7	
β_1	-0.0610	(-1.2853)			0.0423	(0.9094)
β_2	0.0609	(2.1389)			0.0757	(2.3089)
γ_1			-0.3729	(-4.8580)	-0.4625	(-4.4671)
γ_2			-0.0341	(-0.3176)	-0.1894	(-1.6464)
R^2	0.0032		0.0057		0.0068	
Wald	22.2084		38.4560		56.2575	

The first two regressions (Equations 5 and 6) specify the individual volume and liquidity terms respectively, while the last (Equation 7) is the joint regression. The *t*-statistics for the coefficient estimates are provided in parenthesis beside the estimate. The volume and liquidity coefficients for each quadrant are exhibited.

significance is gauged on the basis of a single tailed Student's t-statistic test. Since the volume and liquidity interactions involve return, and since liquidity is a function of volume, we obtain variance inflation factors and verify regressor matrix conditions, to rule out the possibility of multicollinearity. The largest condition number is 6.13 and the largest inflation factor, 4.25, which are both far below the benchmarks of 20 and 10 respectively for multicollinear data[12].

The high proportion of liquidity coefficients that are negative and significant make clear the role of illiquidity in lowering autocorrelation. The next observation that stands out in relief and immediately interests us is that the incidence of negative *and* significantly negative volume coefficients is never as high as the incidence of similar liquidity coefficients in the appropriate corresponding regressions. In addition, we also find that the coefficient of determination of an individual regression that employs a liquidity term is on average almost twice as high as that for an individual regression that relies on an ersatz volume term. We have provided direct evidence that autocorrelation is influenced by the state of liquidity of the market. Such a relationship is perhaps reflected in traded volume too, which in any case, has a weaker effect on autocorrelation.

Then, we observe that the terms associated with the left returns quadrant have a more pronounced effect of dampening autocorrelation than the terms associated with the right returns quadrant. The volume and liquidity coefficients for the left return quadrant (β_1 and γ_1, respectively) do not assume either negative, or significantly negative values with a lower incidence than the corresponding coefficients in the right (β_2 and γ_2, respectively), for any of the regressions in any sample. This is a clue that illiquidity acts to stronger effect when contemporaneous return is negative. To further investigate, we evaluate the probability that the true left quadrant coefficients are lower than the true right quadrant ones. The details of this calculation are provided in the appendix, as are the probabilities for each stock in the entire sample. We find that, though the probabilities that $\beta_1 < \beta_2$ and $\gamma_1 < \gamma_2$ are not large for many of the stocks, the median values for each sample suggest that there is a higher than average chance that these conditions are met. Yet, this is obviously not enough testimony to conclude that the liquidity or volume to act to a greater extent in the lower tail. Therefore, we merely note here that illiquidity tends to act more adversely on serial correlation

[12]The medians of the two measures are only 4.93 and 1.59, respectively.

Table 5: Regression Results for the Sample

Regression	5		6		7	
β_1	0.8710	(0.2258)			0.3226	(0.0645)
β_2	0.2581	(0.0000)			0.2581	(0.0000)
γ_1			0.9677	(0.6774)	0.9677	(0.7097)
γ_2			0.7097	(0.3548)	0.8065	(0.4194)
Sample: DJIA and traded index fund, n=31						
β_1	0.6563	(0.2500)			0.5313	(0.0313)
β_2	0.2813	(0.0313)			0.2500	(0.0625)
γ_1			0.7813	(0.4375)	0.8125	(0.4688)
γ_2			0.5625	(0.1875)	0.6875	(0.1875)
Sample: S&P 500, n=32						
β_1	0.7500	(0.3125)			0.2500	(0.0000)
β_2	0.1875	(0.0625)			0.1250	(0.0000)
γ_1			0.9375	(0.6875)	1.0000	(0.8750)
γ_2			0.8125	(0.4375)	0.8125	(0.5000)
Sample: Individual Stocks, n=16						

As with the last table, the three pairs of columns report results for the volume and liquidity individual regressions and the joint regression, in sequence. The proportion of time series for which each coefficient is negative, as well as the proportion for which each coefficient is negative *and* significant (in parenthesis) is presented. Again, the volume and liquidity coefficients for each quadrant are exhibited.

in the left returns quadrant than in the right.

It is remarkable that the majority of stocks (at least 74% in each sample group) do not show a negative volume effect on autocorrelation for positive returns ($\beta_2 < 0$), since this abets the earlier finding of a relative persistence in return with rising volume [25]. In this investigation, the relationship between autocorrelation and volume turnover over two ranges was sought for an equally weighted index of large firms that tracked the Dow Jones Industrial Average. For positive returns, a clearly positive relationship between autocorrelation and more local increases in volume turnover were discovered simultaneously with a weakly negative relationship between autocorrelation and the longer range volume measure. While these relationships are retained for negative returns, their relative strengths are inverted, so that the negative effect is more pronounced overall. The notion that it is easier to follow trends in a rising market (in regard to return and volume) corresponds to this dominant negative effect of volume (standing in for liquidity) on autocorrelation.

To verify how stable this liquidity relationship is across time, we divide the sample period into approximately four decades, and estimate the same set of eight coefficients as above. We pool all the three sample groups together, but since the stocks in the data set have not all been trading for the same length of time, the number of time series increases with each consecutive decade. Each decade long period yields the same results as the entire period of forty years; these results are weakest in the fourth decade (1993-2004) with sixty stocks. There are two likely reasons for this: this was the decade of the technology boom, when trades that generated extreme returns were more likely to be driven by changes in market valuation of price than by liquidity trades; then, a number of the stocks in this sample have only been trading for about five years or so, and consequently these series are not long enough to betray the same attributes as more complete series.

Before closing this section, we reiterate the effect of market liquidity conditions on autocorrelation in the tails. Since liquidity is lower in the tails than elsewhere, and the state of liquidity (or the lack thereof) has been found to hamper autocorrelation, we associate lower tail autocorrelation with a paucity of market liquidity. In short, we ascribe the anomaly in tail return autocorrelations to prevalent market liquidity conditions.

4 Conclusion

We summarise our findings in this section, then draw conclusions, and from their implications offer some avenues for further investigation.

First, this essay identifies anomalous behaviour in tail autocorrelations: they are lower, when common models of stock returns would have them be larger, than equally heavy central quantile and unconditional autocorrelations. Not only this, but the sign of autocorrelation in the lower tail exhibits a tendency to oppose that of unconditional autocorrelation, which is again contrary to standard predictions. Second, this essay directly confirms the negative effect of market liquidity on autocorrelations: markets in lower liquidity states show weaker autocorrelations than do more liquid markets; this effect has a predilection to be more aggravated for negative returns than for positive returns, so that it is experienced more strongly in the left quadrant than in the right. Our tests of this relationship rely on a direct measure of market liquidity, the Amihud Illiquidity Index, and show that this traces a far clearer causation to dampened autocorrelations than volume turnover, which has been taken to indicate liquidity demand in the market. Noting that liquidity conditions are relatively distressed in the tail quantiles, we explain the anomalous autocorrelation phenomena earlier identified.

As a practical contribution, then, we offer market liquidity as a determinant of either a reversal or a weakening of runs in current trends. A majority of the individual stocks tested here have shown negative serial correlation conditional on extreme negative returns. Though this phenomenon has been reported as early as a decade and a half before, its persistence today might imply existence of scope for its exploitation by speculators. On the whole, positive profits should accrue to those who play the role of market maker when the market is deficient in liquidity. This may have significance for scholars of market microstructure and market design, for it implies that the specialist function can be performed in a decentralised and non-specific manner by profit-motivated market participants. As importantly, this is consequential for investment analysts and anyone who make portfolio allocation choices. It demonstrates the efficacy of the illiquidity measure and the importance of liquidity in profitable equity trading.

Again, our findings reverberate in a general way with the research on market overreaction: when excess returns are realised, it is likely that this is in overreaction as proven by the adjustment over the next few days, or lower conditional returns autocorrelations. Further, this overreaction is more evident in the lower tails than the upper tails, a sign that bad news has more profligate impact than does good news, on return[13].

The results of this work could be taken forward by finding the reasons behind the asymmetric autocorrelation dampening in each tail. It might be that liquidity suffers more on days of extreme negative return than on days of extreme positive return, which would explain how a liquidity effect homogeneous across the range of returns could result in this feature. Another important matter that we have not attempted to resolve here is whether extreme returns are driven by non-informational trades or changes in market valuation; in accordance with the literature we have maintained that this is so, but if it were not, while the relationship between autocorrelation and liquidity would still be upheld, the reasons behind it would need to shift to a more endogenous base. A relevant and important future objective would be to devise a methodology to isolate the liquidity effect from the Roll effect in the bid-ask spread. Not only would this facilitate ready employment of the bid-ask spread as an alternative liquidity measure against which to verify the results of this work, but also it would permit expansion of the set of stocks for which this phenomenon may be investigated. As a first attempt at studying tail autocorrelations and identifying and explaining aberrations in their behaviour, there are doubtless many refinements that would further advance our results. Of these, we mention two: obtaining an expression for the variance in the tails of a process with time-conditional variance would provide confirmation of our results for conditional autocorrelation without resort to the simulations for this; and second, identifying structural changes in liquidity and volume turnover over the years would remove any possible adverse effects this may have on our regressions.

The facts unveiled here, and their conclusions seem to qualify the well-known market precept that it is easier to follow trends in a bull market than a bear market: it is increasingly difficult to do so as the bull market exhibits increasing volatility, or as markets deliver extreme positive returns, in other words.

[13] Of course, these statements are not completely consistent with the standard presumption that extreme returns arise due to settlement of liquidity demands and not propagation of public information, and changes in market valuation.

A Some Details Regarding the Data Set

A.1 Summary Description

Table 6: Sample Descriptive Statistics

	DJIA	**S&P 500**	**Other Stocks**	**Traded Index Fund**
Mean Return	0.0004	0.0003	0.0004	0.0005
	0.0015	0.0034	0.0007	0.0005
St. Dev.	0.0132	0.0152	0.0120	0.0112
	0.0287	0.0514	0.0205	0.0112
Skewness	-0.0004	0.1062	-0.0083	-0.0181
Excess Kurtosis	7.1617	4.8682	15.5149	3.5030
Tenth Decile	-0.0188	-0.0243	-0.0167	-0.0123
Ninetieth Decile	0.0210	0.0279	0.0187	0.0129
Volume Turnover	0.0000	0.0000	0.0000	0.0000
	0.1410	0.7383	0.4556	2.4465
Illiquidity Index	-3.6230	-8.9035	-5.6044	-0.6135
	1.9399	2.7447	2.0471	0.3507
Autocorrelation	0.0301	0.0321	0.0159	-0.0386

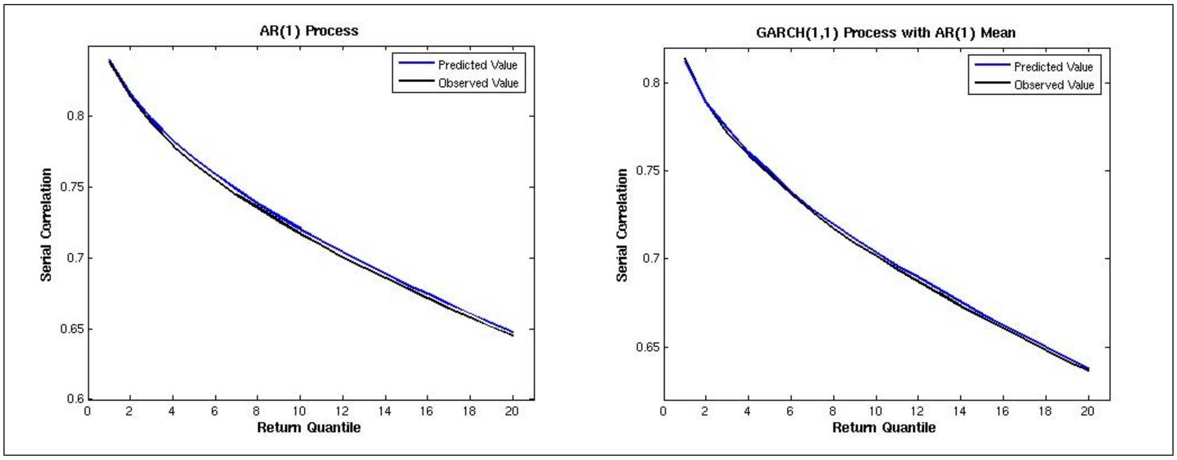
Ranges are provided within each data group for mean return, standard deviation, volume turnover and illiquidity; for the other attributes the median within each group is provided.

A.2 Individual Stocks in the Data Set

Sample I: Alcoa, America International Group, American Express, Boeing Aerospace, Citigroup, Caterpillar, Dupont, Disney, General Electric, General Motors, Home Depot, Honeywell, Hewlett Packard, IBM, Intel, Johnson & Johnson, JP Morgan Chase, Coca Cola, McDonald's, 3M, Altria Group, Merck, Microsoft, Pfizer, Procter & Gamble, AT & T, United Technologies, Verizon, Walmart, ExxonMobil, Spider S & P 500 Traded Index Units (SPDR). **Sample II:** Apple, Abbott Laboratories, Amgen, Bank of America, Comcast, Conoco Phillips, Cisco, Chevron, Dell, EBay, Sprint Nextel Corporation, Goldman Sachs, Eli Lilly, Lowe's Companies, Medtronic Inc., Merrill Lynch, Motorola, Morgan Stanley, Oracle, Pepsi, Qualcomm, Schlumberger, Time Warner Inc., Texas Instruments Inc., Tyco International, United Health Group, United Parcel Service, US Bancorp, Wachovia Corporation, Wells Fargo & Co., Wyeth, Yahoo. **Sample III:** BP Amoco, Commonwealth Edison Co., Dow Chemical Co., Eastman Kodak, Ford, GTE, Imperial Oil Ltd., ITT Industries, Pacific Gas & Electric Co., RJR Nabisco, Southern California Edison Co., Sears Holdings, Southern, Texaco, United States Steel Corporation, Westinghouse.

B Additional Figures

Figure B.1: Forecasts of Conditional Autocorrelation for the AR(1) and AR(1) GARCH(1,1) processes



The processes each have an unconditional autocorrelation coefficient of 0.5 while the errors for the GARCH(1,1) process are specified by $\epsilon_t = h_t^{0.5} z_t$ where $h_t = 0.90 + 0.05\epsilon_{t-1}^2 + 0.70h_{t-1}$ and $z_t \sim N(0, 1)$. The observations are made on series one hundred thousand points in length.

Figure B.2: Probabilities of Lower Pooled Tail Autocorrelations

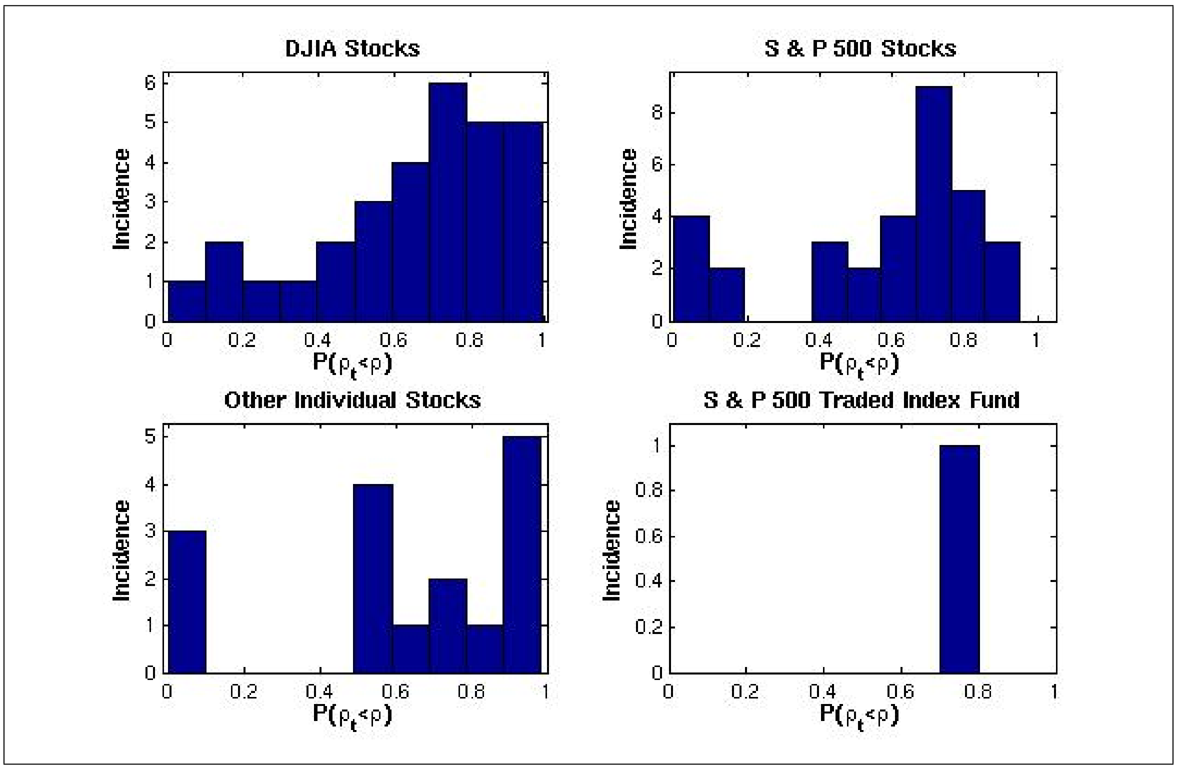


Figure B.3: Probabilities of Lower Tail Autocorrelations being lower than Upper Tail Autocorrelations

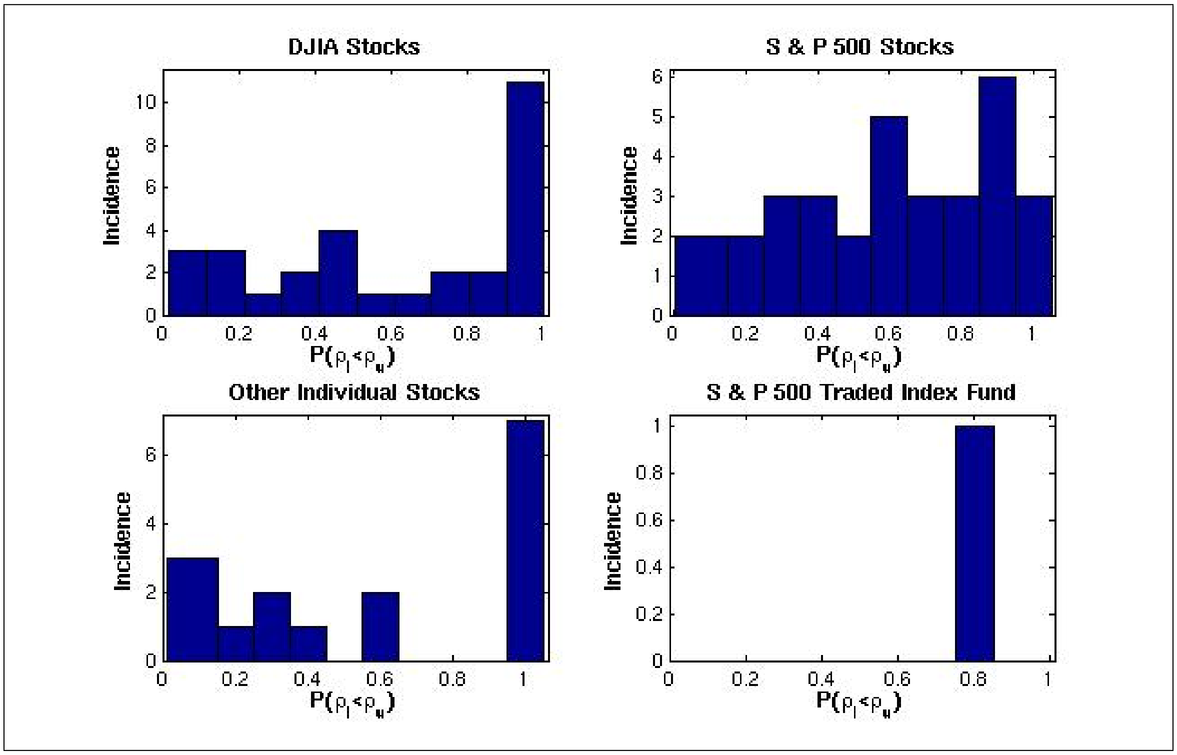
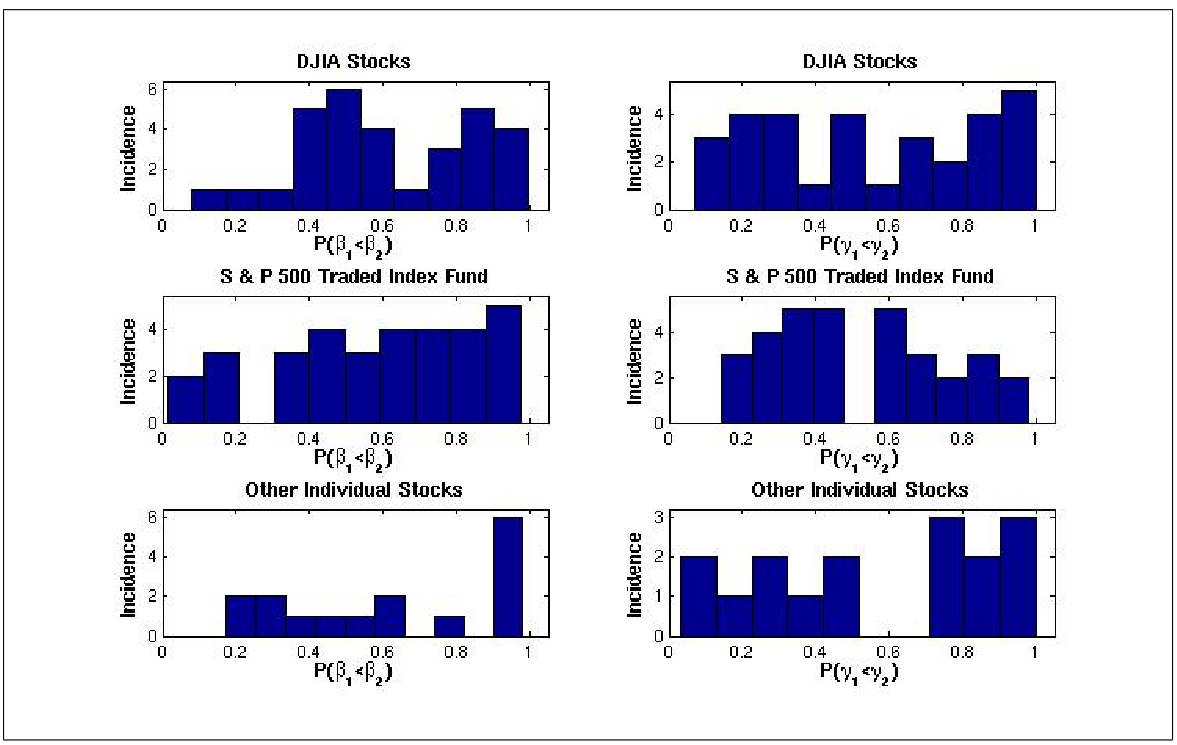


Figure B.4: Probabilities of Disparate Liquidity Effects in The Tails



C Additional Tables

D Estimate Comparisons at Highest Possible Levels of Confidence

In general, if we have two estimates, $\hat{x}$ and $\hat{y}$ of quantities x and y, surmised to be related so that $x < y$, what level of confidence may we use to make this assertion? Our objective is to determine the value of $\Pr(x < y)$ in the inequality $\Pr(x < y) \geq (1-\alpha)(1-\beta)$, which holds when $\Pr(x < x_u) = 1-\alpha$ and $\Pr(y > y_l) = 1-\beta$, where x_u and y_l are one-tailed confidence bounds for the variables. We achieve this by an optimisation programme over the domains α and β, which we subject to the constraint $x_u = y_l$ so that the value, so that the value of $\Pr(x < y)$ may be exactly determined. Here, the problem of comparing magnitudes of correlations has been worked out, but this method may be used for any estimates, in fact.

D.1 Coefficients of Correlation from Two Different Samples

For estimates of correlation between two different pairs of variables with N_x and N_y observations each, $\hat{x}$ and $\hat{y}$ respectively, construct confidence intervals $(-\infty, x_{u,\alpha})$ and $(y_{l,\beta}, \infty)$ at confidence levels of $1-\alpha$ and $1-\beta$, possibly unequal. Then the probability that the population correlation coefficient x is lower than the population correlation coefficient y, $\Pr(x < y) \geq (1-\alpha)(1-\beta)$, with the inequality binding when the upper bound on the smaller hypothesised of the coefficients is equal to the lower bound on the other, or, $x_{u,\alpha} = y_{l,\beta} = q$. Subject to this constraint, our objective is to minimise the value of $\alpha + \beta - \alpha\beta$, or in other words to estimate the highest level of confidence with which x may be said to be lower than y. The confidence bounds for the correlation coefficient are obtained using the Fisher z-transform which is distributed normally with variance $(n-3)^{-1}$ where n is the number of observations in the sample. We may now simplify the constraint to Equation D.3, as shown below.

$$x_{u,\alpha} = \frac{\exp(2z_u - 1)}{\exp(2z_u + 1)}, \text{ where } z_u = \frac{1}{2}\log\frac{1+\hat{x}}{1-\hat{x}} + (\frac{1}{N_x - 3})^{\frac{1}{2}}\phi^{-1}(1-\alpha) \tag{D.1}$$

$$y_{l,\beta} = \frac{\exp(2z_l - 1)}{\exp(2z_l + 1)}, \text{ where } z_l = \frac{1}{2}\log\frac{1+\hat{y}}{1-\hat{y}} - (\frac{1}{N_y - 3})^{\frac{1}{2}}\phi^{-1}(1-\beta) \tag{D.2}$$

$$z_u = z_l, \text{ or, } \beta = 1 - \phi\Big(\frac{1}{2}(N_y - 3)^{\frac{1}{2}}\log\frac{(1+\hat{y})(1-\hat{x})}{(1-\hat{y})(1+\hat{x})} - (\frac{N_y - 3}{N_x - 3})^{\frac{1}{2}}\phi^{-1}(1-\alpha)\Big) \tag{D.3}$$

We may say that correlation coefficient x is lower than y with an utmost confidence level of $1 - \alpha - \beta + \alpha\beta$, where α and β are easily reckoned from the optimisation programme so drafted.

References

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

[2] A. Atkins and E. Dyl. Price reversals, bid-ask spreads, and market efficiency. *The Journal of Financial and Quantitative Analysis*, 25(4):535–547, December 1990.

[3] D. Avramov, T. Chordia, and A. Goyal. Liquidity and autocorrelations in individual stock returns. May 2005.

[4] T. Bollerslev, R.Y. Chou, and K.F. Kroner. ARCH modeling in finance. *Journal of Econometrics*, 52:5–59, 1992.

[5] B. Boyer, M. Gibson, and M. Loretan. Pitfalls in tests for changes in correlations. *International Finance Discussion Papers*, (597), March 1999.

[6] M. Bremer, T. Hiraki, and R. Sweeney. Predictable patterns after large stock price changes on the Tokyo stock exchange. *The Journal of Financial and Quantitative Analysis*, 32(3):345–365, September 1997.

[7] M. Bremer and R. Sweeney. The reversal of large stock-price decreases. *The Journal of Finance*, XLVI(2):747–754, June 1991.

[8] K. Brown, V. Harlow, and S. Tinic. Risk aversion, uncertainty information and market efficiency. *Journal of Financial Economics*, 22:355–385, 1988.

[9] J. Campbell, S. Grossman, and J. Wang. Trading volume and serial correlation in stock returns. *The Quarterly Journal of Economics*, pages 905–939, November 1993.

[10] P. Clark. A subordinated stochastic process model with finite variances for speculative prices. *Econometrica*, 41:135–156, January 1973.

[11] E. Comiskey, R. Walkling, and M. Weeks. Dispersion of expectations and trading volume. *Georgia Institute of Technology Working Paper*, October 1984.

[12] J. Conrad, A. Hameed, and C. Niden. Volume and autocovariances in short-horizon individual security returns. *The Journal of Finance*, XLIX(4):1305–1329, September 1994.

[13] D. Cox and D. Peterson. Stock returns following large one-day declines: Evidence on short-term reversals and longer-term performance. *The Journal of Finance*, 49(1):255–267, March 1994.

[14] G. Duffee. Trading volume and return reversals. *Finance and Economics Discussion Series*, (192), 1992.

[15] T. Epps and M. Epps. The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis. *Econometrica*, 44:305–321, March 1976.

[16] L. Glosten and P. Milgrom. Bid, ask, and transaction prices in a specialist market with heterogenously informed traders. *Journal of Financial Economics*, 13:71–100, 1985.

[17] S. Grossman and M. Miller. Liquidity and market structure. *The Journal of Finance*, XLIII:617–633, 1988.

[18] F. Hamelink. Systematic patterns before and after large price changes: Evidence from high frequency data from the Paris bourse. *Journal of Forecasting*, 22(6-7):533–549, September-November 2003.

[19] L. Harris. Transactions data tests of the mixture of distributions hypothesis. *University of Southern California Working Paper*, April 1984.

[20] L. Harris. Cross-security tests of the mixture of distributions hypothesis. *Journal of Financial and Quantitative Analysis*, 21:39–46, March 1986.

[21] J. Hasbrouck. Trading costs and returns for us equities: The evidence from daily data. *NYU Stern School Department of Finance Working Paper*, February 2005.

[22] J.M. Karpoff. The relation between price changes and trading volume: A survey. *The Journal of Financial and Quantitative Analysis*, 22(1):109–126, March 1987.

[23] A.S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1336, November 1985.

[24] B. LeBaron. Persistence of the Dow Jones index on rising volume. *University of Wisconsin Madison Working Paper*, 1992.

[25] B. LeBaron. Some relations between volatility and serial correlations in stock market returns. *The Journal of Business*, 65(2):199–219, April 1992.

[26] B. Lehmann. Fads, martingales, and market efficiency. *The Quarterly Journal of Economics*, 105(1):1–28, February 1990.

[27] I. Morgan. Stock prices and heteroscedasticity. *Journal of Business*, 49:496–508, October 1976.

[28] D. Morse. Asymmetrical information in securities markets and trading volume. *Journal of Financial and Quantitative Analysis*, XV:1129–1146, 1980.

[29] J. Park. A market microstructure explanation for predictable variations in stock returns following large price changes. *The Journal of Financial and Quantitative Analysis*, 30(2):241–256, June 1995.

[30] E. Sentana and S. Wadhwani. Feedback traders and stock return autocorrelations: Evidence from a century of daily data. *Economic Journal*, CII:415–425, 1992.

[31] R. Westerfield. The distribution of common stock price changes: An application of transactions time and subordinated stochastic models. *Journal of Financial and Quantitative Analysis*, 12:743–765, December 1977.

[32] R. Wood, T. McInish, and J. Ord. An investigation of transactions data for NYSE stocks. *The Journal of Finance*, 60:723–739, July 1985.

[33] C. Ying. Stock market prices and volumes of sales. *Econometrica*, 34:676–686, July 1966.