Persistence of the Dow Jones Index on Rising Volume

Blake LeBaron

Department of Economics

University of Wisconsin - Madison

(608) 263-2516

BLAKEL@vms.macc.wisc.edu

December 1991

Revised: July 1992

ABSTRACT

This paper documents a relation between the persistence of stock returns for a large firm index and trading volume. Previous results on the negative relation between volume and persistence are replicated, but a second effect is discovered. Persistence is directly related to the current rate of change of volume. Also, this effect appears much stronger for positive returns than negative returns. Various specifications are tested to explore the structure of this phenomenon. Finally, individual firm returns are used showing that much of the correlation is coming from cross firm effects involving leads and lags. Some weak evidence is presented showing that lower beta firms are more likely to lead the overall index movements.

Acknowledgments

The author is grateful to Rochelle Antoniewicz and W. A. Brock for helpful comments. This research was partially supported by the Economics Research Program at the Santa Fe Institute which is funded by grants from Citicorp/Citibank, the Russell Sage Foundation, the Alex. C. Walker Educational and Charitable Foundation - Pittsburgh National Bank, and by grants to SFI from the John D. and Catherine T. MacArthur Foundation, the National Science Foundation (PHY-8714918) and the U.S. Department of Energy (ER-FG05-88ER25054). The author also is grateful to the National Science Foundation (SES-9109671) for support.

I. Introduction

For many years the weak short range correlations in index stock returns has been thought to be well understood. The explanation given in Fisher(1966) which stressed the delayed adjustment of some firms' prices to new information worked very well in explaining observed empirical results. It could account not only for correlations in broad indices, but also the observed cross correlations from small to large stocks. Recently, this explanation for short range correlations has come under some criticism. Several papers have suggested that the magnitude is too large for certain specific models of nontrading and stale prices.[1] Others have noted a changing structure in the correlation pattern.[2] They generally find an inverse relation between correlations and some measure of trading activity, either volatility or trading volume.

One important aspect of these results on changing correlations is that they are all still qualitatively consistent with Fisher's explanation. However, the mechanism must now be defined in a more elaborate framework. To do this we need the idea of changing speeds of economic time put forth in papers by Clark(1973) and Stock(1987), and market micro-structure frictions summarized in Cohen et al.(1980,1986). These two ideas can be put together to provide explanations which are qualitatively consistent with the observed correlations. During periods when there is very little economic activity, and information is moving slowly, trading frictions would loom large. There may be a large number of transactions carrying over on specialist books, and transactions costs themselves would be large relative to the magnitude of incoming information. The combination of these effects may cause persistence in observed prices and indices. However, during periods of greater activity economic time is moving faster and the daily observations are actually spaced farther apart in economic time. Transactions costs and other frictions will be small relative to the magnitude of events. Observed persistence in indices and individual stocks during these periods would be small.

Initially it appears that this is a useful explanation, and all that is necessary is to begin exploration into quantifying it and calibrating models. However, there are still several troubling facts that directly contradict some of the previously stated results. First, in an early paper, Morse(1980) demonstrates evidence for

[1] For examples of this see Atchison et al. (1987) and Lo and Mackinlay(1990). Also, Lo and Mackinlay(1990) and Mech(1991) analyze lagged price adjustment issues.

[2] These papers include Campbell etet al.(1991), LeBaron(1992), and Sentana and Wadhwani(1990) for stocks. Bilson(1990), Kim(1989), and LeBaron(1992) for foreign exchange. A very detailed study on individual firms is Wiggins(1991).

increased persistence in individual firms on higher volume. He explains this by appealing to an idea of diverse information. Periods of high volume will be periods when traders' information is more diverse and is in the process of converging. During these periods learning will be taking place and beliefs will be converging. Correlations will be induced by this convergence mechanism. In more recent work Antoniewicz(1992) documents increases in correlation on higher trading volume for individual NASDAQ firms which are consistent with Morse(1980)'s results, indicating more persistence when trading volume is high. They also agree with a large bulk of technical trading rules which suggest that traders should follow price trends when they occur on generally large volume.[3]

This paper will address problems suggested by these very different results. Evidence for increased persistence for the Dow Jones Index on high volume will be considered. Many of the results stated earlier were applied at the individual firm level to cross sections of firms of varying sizes. Most of the analysis here is done using conditional correlations estimated using aggregate volume conditioning information.

Section II describes the data series used. Section III presents the important empirical message of the paper in some simple plots of conditional correlations and some initial parameter estimates. Section IV tests these results more rigorously by running several specification tests for appropriate functional form. Section V addresses index and portfolio issues as the causes of these effects. Section VI tests a simple trading strategy using volume information. Finally, section VII concludes the paper, and connects some of these results to recent theories of trading volume.

II. Data Descriptions

This study concentrates on an equal weighted index of large firms designed to closely track the Dow Jones index. The list of the 21 firms in this index is given in table 1. Many of these firms were members of the Dow for much of the sample period. Several are firms that have been added, or have left. Daily returns for these firms are collected from the Center for Research in Securities Prices (CRSP) and an index is created from an equal weighted portfolio. The objective of using this index is to get a group of large homogeneous firms where the problems of transactions costs, nontrading, and private information are minimized. The Dow Jones index serves this purpose well. Constructing a pseudo Dow index from scratch allows more direct control over how the index is handled and the inclusion of dividends in the returns series. Most importantly,

[3] Also, one of the tests performed in Wiggins(1991) finds some evidence for increased cross firm correlations when this correlation is conditioned on contemporaneous volume increases.

it allows tests to be performed on disaggregated components of the index. The Value Weighted index from CRSP will also be used for comparisons with previous results and CAPM beta estimation. The time period covered begins in July 1962 and extends through September 1987.

The second important series used is trading volume. For this the total NYSE shares traded will be used.[4] The raw trading volume will have to be processed in several ways. First, the series is converted to turnover ratios by dividing by the number of shares outstanding from the Survey of Current Business published by the U.S. department of Commerce.[5] Turnover ratios are plotted in figure 1. The initial difficulty in using this series is that it is nonstationary. However, the nonstationarity appears to come in several discrete jumps as opposed to a prolonged trend over the entire sample. As an initial detrend procedure a linear trend is fit to the log of the raw turnover series and the residuals of this are plotted in figure 2. This eliminates most of the trend, but some very long range movements still appear to exist in the series. A second detrending procedure divides the raw series by a 100 day moving average and then logs this ratio. This is shown in figure 3 which clearly displays a better transformation of the series from the standpoint of removing long range trends.

Some care should be taken in interpreting the results of a moving average detrending procedure since the moving average may induce spurious patterns on the transformed series. Since this paper will use volume as an exogenous information variable and is not concerned with its dynamics per se this will not be an important problem.[6]

A second volume series is generated following a similar procedure, but replacing the 100 day moving average with a 5 day. This series is designed to pick up short range movements in trading volume relative to the past several days. Initially, the log difference of the volume series was tried, but this series turned out to be too noisy. Similar results have been obtained for several other short range MA's, but 5 days was used to minimize daily seasonals in the MA.

Finally, daily seasonals are removed from both volume series. Results of this linear regression are given in table 2. The series detrended by the 100 day moving average will be referred to as the volume (v_t)

---

4 This series is available for a much longer time period in Pierce(1991).

5 See Mulherin and Gerety(1988) for some detailed information on the properties of this series.

6 An early paper on this subject is Granger and Hughes(1971). These authors found a proportionate MA detrend did not change the important structures found in the Beveridge wheat series.

series, and the five day moving average detrended series will be referred to as (dv_t). These series are the final residuals after daily seasonals have been removed.

Summary statistics are given in table 3. The index formed from the 21 firms in table 1 is labeled Dow21. For the Dow series we see the usual amount of large kurtosis present in financial series. The first order correlations show significant correlation at the first lag and no correlation after that. This is typical for a large firm index. The magnitude of these numbers is also crucial since much of this study will be concerned with correlations at 1 lag conditioned on trading volume.

The transformed volume series, v, looks closer to normal than the returns series with a kurtosis of 3.7. Also, taking logs has eliminated much of the skewness in the series. Trading volume is highly persistent here as seen in the large autocorrelations. The second volume series, dv, looks much less normal with a kurtosis of 7.3. Also, the autocorrelations reflect the importance of the short range MA on autocorrelation with a large drop to negative correlations after a positive correlation at lag 1. Figure 4 shows 20 lags of autocorrelations for both volume series along with the 95% Bartlett bands. This figure clearly shows the dramatic persistence of the v series and the unusual behavior of the dv series induced by the short range moving average. Figures 5 and 5b display cross-correlations of volume with returns and absolute values of returns respectively.[7] These figures show a strong contemporaneous correlation between volume both volume series and the return series. There is little connection between volume and either return series in the future.

III. Conditional Correlations

This section presents some pictures examining the movements of conditional autocorrelations in the series. Autocorrelations in returns will be estimated over different ranges of the two volume series v_t and dv_t. Figure 6 plots estimates of the correlations for the Dow series from July 1962 through September 1987. The estimated correlation from return at time t to $t+1$ is plotted conditional on time t volume information. The series is unconditionally demeaned before the correlations are estimated.[8] Both volume and the change in volume are mapped into their distribution fractiles before the correlations are estimated. Then a nonparametric estimation of the correlation is performed using a uniform kernel and a bandwidth of

[7] Many of the well known connections between returns and volume are summarized in Karpov(1987).

[8] Similar results were found using means estimated within each volume grouping

0.3. This can be viewed simply as moving a $0.3x0.3$ square box around on the volume fractile base and plotting the estimated correlation in the vertical direction.

This figure is very informative. Moving along the volume axis on the right, the effect of increasing volume can be observed. Moving from low to high (back to front) on this axis the general decrease in correlation identified in Campbell, Grossman, and Wang(1991) (CGW) can be seen in the decrease in correlations as volume increases. Moving from left to right along the front axis a second effect is observed. Conditioning on the local increase in volume there is actually an increase in correlation as we move from the low to high (right to left). This increase appears to flatten out at higher levels of volume. This effect shows some increased persistence in the Dow on a local increase in volume as well as the original decrease in correlation with higher overall volume.

This appears to start to agree with some of the results in Morse(1980) and Antoniewicz(1992) on volume and persistence. Separating the returns into positive and negative returns at time t makes this result more intriguing. Figure 7 plots the relation for positive returns at time t.[9] The increasing delta volume effect is strengthened while the volume effect is greatly reduced. Figure 8 plots the same results for negative returns. Here, we see a reverse picture with a decrease in the delta volume effect and an increase in the volume effect. These results are broadly consistent with some of the asymmetries pointed out by technical analysts. They spend many pages talking about the use of volume confirmation in rising markets, but few pages on what should happen in falling markets.[10]

Also, it is possible that the results for negative returns are related to some of the reversal phenomena documented in Bremer and Sweeney(1991). These authors show that large price decreases tend to be followed by a price increase. Examining figure 8 more closely shows that the delta volume effect appears to be present at lower levels of volume, but disappears at higher levels. This is also consistent with the theory of investor behavior suggested in Brown et al.(1988) where investors are faced with news shocks for which the uncertainty is resolved over several days. During this resolution period prices will be rising for both positive shocks and negative shocks. This generates persistence for positive shocks and reversals for negative shocks.

[9] In this case r_t uses the mean over only the positive returns for the calculation of the correlation with r_{t+1}.

[10] See Weinstein(1988) page 237. In his section on short selling advice very little emphasis is placed on trading volume.

The results presented in these plots are suggestive of what is going on, but they should not be viewed as statistical tests. There are still many problems that are not being accounted for in these pictures. There is an obvious dependence between the two measures so the box sizes are not uniform. Also, the conditional means are not adjusted for. Finally, the impact of outliers needs to be accounted for. Some of these issues will be addressed in the next section where the results will be made more precise.

The results presented in figures 6 through 8 are directly tested in table 4. The effect of both v_t and dv_t on the correlation pattern of returns is estimated. The first row of table 4 shows the estimates from fitting an unconditional AR(1) to the returns series.[11] As was evident from table 3, this term is significantly different from zero. The results in CGW showed a negative relation between volume at t and the conditional correlation between returns at time t and t+1. This result is repeated for the Dow in the next row of table 4. The linear volume correlation term, β_1 is significantly negative. In the bottom section of this table the results are presented for the CRSP value weighted index. The estimated parameters and reported R^2 are very close to those in CGW which is interesting since the volume series is detrended slightly differently.

The row labeled dv in table 4 adds the delta volume effect shown in the figures. A term is added to the conditional correlation adding the dv_t series to the conditional correlation expression for volume for positive returns only. The parameter β_2 is used for this term. Its estimated value is significantly positive as predicted by the earlier pictures. Also, a parameter is added for negative returns, β_3. Consistent with the figures this parameter is not significantly different from zero. These results suggest that there may be two effects connected to trading volume. One related to its overall level, and a second to its local rate of change. This is seen as an increase in persistence when the Dow is rising, and volume is locally rising.

Figure 9 plots the conditional correlation over all returns for varying v and dv levels using the estimated relation from table 4. This figure is constructed using the central 98% of the v and dv distributions. It gives some idea of the magnitude of the correlation changes from a high of 0.4 to a low of about -0.2.

IV. Specification Tests

There are many problems in confirming that the previous results are clearly indicating persistence on a rise in volume. In this section several different specifications are tested to try to find out the cause.

[11] The results in this table were all estimated using OLS with White heteroskedasticity consistent standard errors in parenthesis.

The first possibility is that volume is a noisy measure of overall economic activity, and this activity is what is related to the changing correlation pattern. This suggests a smoother index to measure activity such as,

$$\frac{ma_t^5}{ma_t^{100}}, \tag{4.1}$$

where ma^n is the moving average of volume over n days. Since the two volume series used here are logged this measure is equivalent to $v_t - dv_t$. The asymmetry over the positive and negative returns immediately suggests that something different is going on here. However, the exact nature of the phenomenon is still not clear. The models estimated imply correlations as a function of volume as,

$$\rho(v_t, dv_t) = \beta v_t + \gamma dv_t^+, \tag{4.2}$$

where dv_t^+ indicates only for positive returns at t. However, this could also be written as,

$$\rho(v_t, dv_t) = \beta v_t + \gamma(dv_t - dv_t^-). \tag{4.3}$$

If $\beta = -\gamma$, then

$$\rho(v_t, dv_t) = \beta(v_t - dv_t) + \beta dv_t^-. \tag{4.4}$$

Therefore, in the fitted linear specification for the estimated parameters it is not clear whether there is an adjustment for the one day volume effect for positive returns, or adjustment to a $v - dv$ index for negative returns using dv.

This is further tested in the following regression,

r_t=	a+	$(\beta_0+$	$\beta_1 S_t v_t+$	$\beta_2 \bar{S}_t v_t+$	$\beta_3 S_t dv_t+$	$\beta_4 \bar{S}_t dv_t) r_{t-1}$
		0.137	-0.322	-0.530	0.343	0.141
		(0.017)	(0.120)	(0.149)	(0.158)	(0.197)

This tests the impact of both the v and dv terms over positive and negative returns at time t. The standard errors are heteroskedasticity consistent standard errors. Testing equality of β_1 and β_2 gives a chi-squared statistic of 1.15 which has a p-value of 0.28. The hypothesis of $\beta_1 = -\beta_3$ is tested giving a chi-squared value of 0.038 with a p-value of 0.84. The inability to reject both these relations suggests that we are in the situation presented above where there are two very different possibilities for what is going on.

We are left with two possible functional forms for the correlation,

$$\rho = f(dv_t - v_t, dv_t^-), \tag{4.5}$$

or,

$$\rho = f(v_t, dv_t^+), \tag{4.6}$$

which given the estimated parameters and linear specification are indistinguishable. A graphical attempt at distinguishing these is given in figure 10. It is a figure similar to 6-8 now using $v_t - dv_t$ as one of the pieces of conditioning information. If the first specification were correct then there would be no adjustment to correlations for dv looking at positive returns. The figure shows some changes exist. It should be compared with figure 8, which if the second specification were true would show no change over dv. Both figures show small changes over dv and are not dramatically different. This problem clearly needs a better statistical test of functional form, but the figures presented show that even a precisely tuned test may have a difficult time determining the correct specification. For the remainder of this paper the second specification above will be considered.

Table 5 presents further tests of this specification. In the first two rows the model is estimated over two subsamples. The sign patterns in the coefficients are consistent over the two subsamples. However, the volume effects appear stronger during the first subsample. The t-statistic for β_2, the dv coefficient, falls to 1.73 during the second subsample which is still significantly positive at the 10% significance level.

The third rows test for the impact of outliers on the volume correlation relation. Both volume series are transformed to their individual fractile rankings less 0.5 (min = -0.5, median = 0, max = 0.5). Then the same regression is run using these transformed series. The results show little difference using these transformed series.

The fourth row checks to see if the dv effect is coming in because of a misspecification in the volume-return relation. The expected return is fit using a nonparametric estimate,

$$r_{t+1} = f(r_t, v_t) + \epsilon_t. \tag{4.7}$$

To test this a kernel density estimate of the expected value of r_{t+1} is fit. A uniform kernel is used with bandwidths chosen using cross validation with a squared deviation loss function. The estimated $f()$ is shown in figure 11. This plot shows the increase in correlation as volume falls. It also shows that a linear specification for the r_t to r_{t+1} is probably not a bad approximation. Residuals of this estimated model are then sent through the estimation in table 5. The significance of the β_2 parameter suggests that the estimated model has not removed much of the information coming from the dv series. However, it should be noted that the model did not remove much of the impact of v either, suggesting tests of other bandwidths and kernels.

The fifth row of table 5 replaces dv with the residual of an AR(5) fitted to the v_t series. This attempts to capture the idea of a surprise component of the volume series. The results are similar to those using the actual dv series suggesting that these two values measure similar components of volume. They are both large when volume is unusually large relative to the recent past. The final row repeats this test replacing dv with the residuals of dv regressed on contemporaneous v. This tests the correlation pattern when the component of dv linearly orthogonal to v is used. The results do not change dramatically.

Table 6 checks to see how far into the future this phenomenon persists. The relation is estimated at lags 2 and 3. The table shows a strong weakening of the results by lag 2. The parameter estimate for β_1 is insignificant for all the tests. However, β_2 is significant at lag 2, but this goes away at lag 3. Also, for all the estimated regressions the R^2's are very small. These results suggest some weak dependence at 2 days in the future, but the effect is gone by day 3.

Table 7 fits a parametric nonlinear model for the changing correlation. The model used is based on the smooth threshold autoregressive model from Tong(1990). This model posits a correlation value which changes as a function of x,

$$a + \frac{b}{1 + e^{cx}}. \tag{4.8}$$

This model allows an smooth change in the correlation as x changes, but the correlation will reach a well defined maximum and minimum when x gets very small, or large. The model as proposed by Tong would have lagged values of the returns series in place of x. Here, the volume series is used in place of x, the variable controlling the level of correlation. Estimated parameters are given in table 7. Estimation was done using nonlinear least squares. The model shows a weakly significant relation for the first nonlinear component, β_1, but all others are clearly not significantly different from zero. Also, the R^2's are not very different from those from the fitted linear specifications. Testing the significance of the nonlinear terms as a group would be difficult since the model contains parameters which are not identified under the null of no nonlinear effects. A very simple test is done to show the lack of importance of the nonlinear effects in this model. In figure 12 the function for the correlation as a function of volume estimated with only the effect from v (first row of table 7) is plotted. The upper figure plots this function over the central 98% of the volume distribution. The function is not dramatically nonlinear over this range. The bottom panel plots the function over a much wider range. This shows that at the estimated parameters the nonlinearity of this function does not dramatically affect the results, indicating a near linear relation, or perhaps misspecification of this functional form.

Table 8 tests whether similar results can be observed outside of the linear regression framework. Specifically, it tests the probability of the same signed return occurring at time t+1 as t for different levels of v and dv. For both v and dv below trend (< 0) the table shows a probability of a positive return at $t + 1$ given a positive return at t of 0.557. However, if $dv > 0$ this increases to 0.605. The number in the table between these two values (3.60) is the t-statistic for equality of the two using the normal approximation for a binomial test, setting the variance, $\sigma^2 = 0.25/N$. It is seen in the table that the impact of dv is greatly reduced when $v > 0$. Also, it is clear that the impact of v is strongest when $dv > 0$. The bottom panel of the table shows very similar results for negative returns. This is somewhat surprising given the asymmetries shown in table 4.

V. Individual and Index Effects

All of the tests performed up to this point have been simply applied to an index of returns without considering the issue of index versus individual returns. This section addresses some of these issues. The results presented here will continue to use the aggregate volume series used in the previous tests. This clearly limits what can be found to volume shocks which are most likely related to aggregate macro shocks. Recent studies by Antoniewitz(1992) and Wiggins(1991) used individual firm volume and are better able to test these types of shocks.

The first row of table 9 fits the model estimated in table 4 to each individual firm separately. The sign series S_t is still set to the sign of the index of the 21 firms, but lagged returns from only each firm individually are used to explain that firm's future returns. A new index is constructed from the residuals of these fitted models, and the model is again estimated over this index. The results given in table 9 are mixed. They show that fitting individual models has no effect on the β_1 affect from v, but they do reduce the significance of the β_2 term from dv.[12] This suggests that much, but not all of the changing correlations are coming from correlations across different firms.

[12] β_2 is still marginally significant with a two-tailed p-value of 0.11.

The second row of table 9 looks at residuals of a fitted market model. For each firm a model of the form,

$$r_{it} = a_i + \beta_i r_{mt} + \epsilon_{it}$$

is fit, and an index is built from the mean of the residual series ϵ_{it}. Beta is estimated using a rolling estimate over the previous 1000 days. The CRSP value weighted index is used as a proxy for r_{mt}. It is clear from the insignificant estimated parameters that removing the market component from the individual firms removes any evidence for the changing correlation. This clearly shows that this is most likely a macro phenomenon, and the correlations are not related to correlations in a few firms moving in ways unrelated to broad market moves.

If index correlations are more likely caused by firms leading and lagging a broad market index then finding out which firms are doing this will be important to understanding why this group of large firms does not move as one unit. Table 10 looks at some portfolios constructed from subsets of the 21 firm index to see if there are any systematic patterns. The first three rows examine portfolios constructed according to market beta's estimated over the previous 1000 days using the CRSP value weighted index. The 21 firms are split into three equal portfolios sorted by beta. The beta portfolios at $t+1$ are regressed on the Dow21 index at t to see if any of these appear to be lagging the index. For β_0 the results are remarkably similar for each of the three portfolios. However, for the other two parameters the effects show a much stronger lagging relation related to volume for the two higher beta portfolios. This gives some evidence that lower beta firms may tend to lead in this index.

A very direct test is performed in the bottom of table 10. The index is split into large moves and small moves on a given day. If on day t the Dow21 index goes up, the large move portfolio is defined as all firms with returns above the median for that day. The small moves are defined as the median and below for that day. This definition is reversed when the market falls. This captures those firms showing the greatest reaction, or driving the index on a given day. This grouping of firms is then followed into the next day. For example, the row labeled "small on small" finds the set of small move firms at time t and regresses the returns for these same firms at $t+1$ on their time t return. If the correlations were traveling across similar groups of firms the parameters should be largest when the two portfolios are constructed from the same groups. However, if the correlations are traveling across different firms in the index then results will probably suggest a leading pattern from large to small movers. The estimated parameters strongly suggest the latter. The regression of small on small shows no significant volume conditioning effect, but it does have a significantly large constant correlation, β_0. The regression of small on large shows a strong lead pattern coming from the volume terms from the large movers to the small.

VI. Simple Trading Rule Results

This section presents some brief results using both volume series for a simple trading rule. The rule is extremely trivial, and the tests are run only to get a general feel for the economic magnitude of some of these findings. The rule will simply buy and hold for a period of one day when returns are high today, and the volume signals indicate a high positive correlation with tomorrow's return. Returns are calculated across the index as a whole.

Results of this experiment are given in table 11. The row labeled "All" shows the overall return for all days. The row labeled $r_t > 0$ shows the return at t+1 conditioned on a positive return at time t. This would be a simple strategy utilizing the unconditional correlations in the series. This strategy generates a return of 0.13 percent per day. Requiring that the return on day t be > 0.004 increases this return to 0.16.[13] Note, that this strategy is in the market only about a third of the time, and generates 2519 trades, constantly changing its position. Requiring that $v < 0$ does not change the conditional return by a large amount. It increases to only 0.17 percent. Finally, requiring also that, $dv > 0$, increases the return to 0.22 percent. The magnitude of these returns is probably not large relative to the transactions costs that they would incur. For even the best traders, roundtrip transactions costs below 0.2 percent are probably difficult to achieve.[14]

Even though the ability of this dynamic strategy to significantly alter returns for a trader appears small, the impact of dv is interesting. The addition of trading volume alone did not have much of an impact, but using dv as well changed the conditional returns dramatically. Only for this test was the return significantly greater than the return not using volume (p-value = 0.06, 1 tailed test). This indicates the importance of dv in a forecasting context.

These tests show that a simple strategy of following the market given volume signals would probably not do well beyond transactions costs. It remains to be seen whether modifications to this rule could change these results. Also, more formal tests will require further parameter tuning and rigorous out of sample tests.

[13] 0.004 is roughly 1 half the standard deviation of the returns series.

[14] See Chan and Lakonishok(1991) for some examples.

This results are only intended as an initial exploration. They do show an interesting effect of using dv as a piece of conditioning information. It appears to be important in helping to increase the conditional return for the trading strategy.

VII. Conclusions

This paper has demonstrated that there is evidence for increased persistence in the Dow Jones Index on rising volume. The volume effect is actually more complicated than previously thought with an inverse relation between the level of trading volume and correlations, and a positive relation between correlation and the local rate of change of volume. This second effect appears to be asymmetric across positive and negative returns. The changing correlation patterns for most of the volume effects are shown to be coming primarily from index cross correlations as opposed to own firm correlations.

While the empirical results here appear strong and consistent across the subperiods they should still be viewed with some caution. In nonlinear modeling it is difficult to test for all possible specifications. Table 5 makes some first attempts at this, but further tests on just how lagged volume should enter into the relation are necessary to sharply demonstrate that it is coming in through the rate of change of volume. Also, the evidence is still inconclusive as to whether adjustments to correlations are occurring on days with positive or negative returns.

These results, showing that the relation between volume and correlations may be quite complex, are related to recent theoretical work on trading volume. Wang(1991) shows that large volume may be associated with more negative or positive autocorrelations depending on whether there is informational asymmetry. Without informational asymmetry large volume indicates a large amount of buying or selling for liquidity reasons and the price should rebound quickly to its previous levels. In the presence of informational asymmetries large volume may be connected with persistence in price movements since the price does not fully reflect the private information of informed traders.

Another interesting theoretical paper which is related to some of the results seen here is Blume, Easley, and O'Hara(1991). This paper introduces trading volume into heterogeneous rational expectations model. Volume can be used by traders to indicate the quality of information signals coming in. This quality is changing over time in the model. Part of this model outlines the connection between the informativeness of the signal and trading volume. For very uninformative signals volume is low since traders place very little confidence in there signals. As the precision of the signal increases, trading volume increases at first.

However, when the informativeness of the signal gets very large trading volume starts to fall due to the fact that people are actually receiving very precise and highly correlated signals. This model also shows how the value of technical analysis depends critically on how informative new signals are for traders relative to prior information. These results again stress a rather complex relation between price movements and trading volume.

The theoretical possibilities are further complicated by results such as those in Kim and Verrecchia(1991) which emphasize the impact of new information arrival, diversity of opinions, and market liquidity. This model suggests that when new information arrives traders diverse interpretations of this information cause more heterogeneity of beliefs in the market, and therefore bid-ask spreads widen and the overall liquidity of the market falls. However, this may be accompanied by increases in trading volume as the informed traders have more diverse beliefs. Uninformed traders stay out of the market at this time. This also suggests several different ways in which volume may interact with price movements depending on what type of volume it is (liquidity or information). It also suggests that the volume volatility connection might be very complicated.

These papers show the complexity possible in volume price relationships, but they do not help much in explaining the asymmetry observed here between positive and negative returns. This could be related to the model of Brown et al.(1988) in which it takes time to resolve the uncertainty connected with new information shocks. Both good and bad shocks have a negative uncertainty impact on prices along with their good or bad impact. For a good shock there is an initial jump and upward persistence after that as uncertainty is resolved. For a bad shock there is an initial downward jump and then a reversal as uncertainty is resolved. This model is consistent with the asymmetries observed here.

Clearly, more theoretical and empirical work is needed in this area. It still is unknown whether we will be able to use volume data for both aggregate and individual stocks to sort out between various competing theoretical models for heterogeneity and learning dynamics. However, the challenge is an exciting one.

References

Antoniewicz, Rochelle. 1992. An Empirical Analysis of Stock Returns and Volume. Ph.D. Disertation. University of Wisconsin, Madison, Wisconsin.

Atchison, M., K. Butler, and R. Simonds. 1987. Nonsynchronous Security Trading and Market Index Autocorrelation. *Journal of Finance* 42 (March) : 111-118.

Bilson, J. F. 1990. "Technical" Currency Trading. In The Currency-Hedging Debate. Edited by L. R. Thomas. London: IFR Publishing Ltd.

Blume, L., D. Easley, and M. O'Hara. 1991. Market Statistics and Technical Analysis: The Role of Volume. Cornell University, Ithaca, NY.

Bremer, M., and R. Sweeney. 1991. The Reversal of Large Stock-Price Decreases. *Journal of Finance* 46 (2) : 747-754.

Brock, W. A., J. Lakonishok, and B. LeBaron. 1992 forthcoming. Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. *Journal of Finance*.

Brown, K. C., W. Van Harlow, and S. M. Tinic. 1988. Risk Aversion, Uncertainty Information, and Market Efficiency. *Journal of Financial Economics* 22 : 355-385.

Campbell, John, Sanford J. Grossman, and Jiang Wang. 1991. Trading Volume and Serial Correlation in Stock Returns. Princeton University, Princeton, NJ.

Chan, Louis K. C., and Josef Lakonishok. 1991. Institutional Trades and Intra-Day Stock Price Behavior. College of Commerce, University of Illinois, Urbana-Champaign, IL.

Clark, Peter K. 1973. A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica* 41 (January) : 135-55.

Cohen, K. J., G. A. Hawawini, S. F. Maier, R. A. Schwartz, and D. K. Whitcomb. 1980. Implications of Microstructure Theory for Empirical Research on Stock Price Behavior. *Journal of Finance* 35 (May) : 249-257.

Cohen, Kalman J., Steven F. Maier, Robert A. Schwartz, and David K. Whitcomb. 1986. *The Microstructure of Securities Markets*. Englewood Cliffs, New Jersey: Prentice-Hall.

Colby, R. W., and T. A. Meyers. 1988. *The Encyclopedia of Technical Market Indicators*. Homewood, IL: Dow-Jones Irwin.

Conrad, Jennifer, and Gautam Kaul. 1989. Mean Reversion in Short-Horizon Expected Returns. *The Review of Financial Studies* 2 (2) : 225-240.

Fisher, L. 1966. Some New Stock-Market Indexes. *Journal of Business* 39 (January) : 191-225.

Gallant, A. R., P. E. Rossi, and G. Tauchen. 1990. Stock Prices and Volume. North Carolina State University. Chapel Hill, NC.

Granger, C. W. J., and A. O. Hughes. 1971. A New Look at Some Old Data: The Beveridge Wheat Price Series. *Journal of the Royal Statistical Society* A 131: 83-99.

Karpov, Jonathan M. 1987. The Relation Between Price Changes and Trading Volume: A Survey. *Journal of Financial and Quantitative Analysis* 22 (March) : 109-126.

Kim, Chong Man. 1989. Volatility Effect on Time Series Behavior of Exchange Rate Changes. Korea Institute for International Economic Policy, Seoul, Korea.

Kim, O., and R. E. Verrecchia. 1991. Market Liquidity and Volume Around Earnings Announcements. UCLA, Los Angeles, CA.

Lakonishok, Josef, and Seymour Smidt. 1988. Are Seasonal Anomalies Real? A Ninety-Year Perspective. *Review of Financial Studies* 1 (4) : 403-425.

LeBaron, B. 1991a. Technical Trading Rules and Regime Shifts in Foreign Exchange. University of Wisconsin, Madison, Madison, WI.

LeBaron, B. 1991b. Transactions Costs and Correlations for a Large Firm Index. Unversity of Wisconsin, Madison, Madison, WI.

LeBaron, Blake. 1992 forthcoming. Improving Forecasts Using a Volatility Index. *Journal of Applied Econometrics*.

LeBaron, B. 1992. Some Relations Between Volatility and Serial Correlations in Stock Market Returns. *Journal of Business* 92 (2) 199-220.

Lo, Andrew W., and A. Craig MacKinlay. 1990. An Econometric Analysis of Nonsynchronous Trading. *Journal of Econometrics* 45 (July/August): 181-211.

Lo, Andrew W., and Craig A. MacKinlay. 1990. When Are Contrarian Profits Due To Stock Market Overreaction? *Review of Financial Studies* 3 (2) : 175-206.

Mech, Timothy S. 1991. What Causes Lagged Price Adjustment? Boston College, Chetnut Hill, MA.

Morse, Dale. 1980. Asymmetrical Information in Securities Markets and Trading Volume. *Journal of Financial and Quantitative Analysis* 40 (5) : 1129-1148.

Mulherin, J. H., Gerety, M. S. 1988. Trading Volume on the NYSE during the Twentieth Century: A Daily and Hourly Analysis. Office of Economics Analysis, SEC.

Pierce, Phyllis S. 1991. *The Dow Jones Averages: 1885-1990*. Homewood, Il: Business One Irwin.

Sentana, Enrique, and Sushil Wadhwani. 1990. Feedback Traders and Stock Return Autocorrelations. London School of Economics, London.

Tong, H. 1990. *Non-linear Time Series: A Dynamical Systems Approach.* Oxford: Oxford University Press.

Stock, James H. 1987. Measuring Business Cycle Time. *Journal of Political Economy* 95 (December) : 1240-61.

Wang, J. 1991. A Model of Competitive Stock Trading Volume. Sloan School, MIT, Boston, MA.

Weinstein, S. 1988. *Secrets for Profiting in Bull and Bear Markets.* Homewood, IL: Dow-Jones Irwin.

White, H. 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test For Heteroskedasticity. *Econometrica* 48 : 817-38.

Wiggins, James B. 1990. Trading Volume and Stock Price Response to Market-Wide Information. Cornell University, Ithaca, NY.

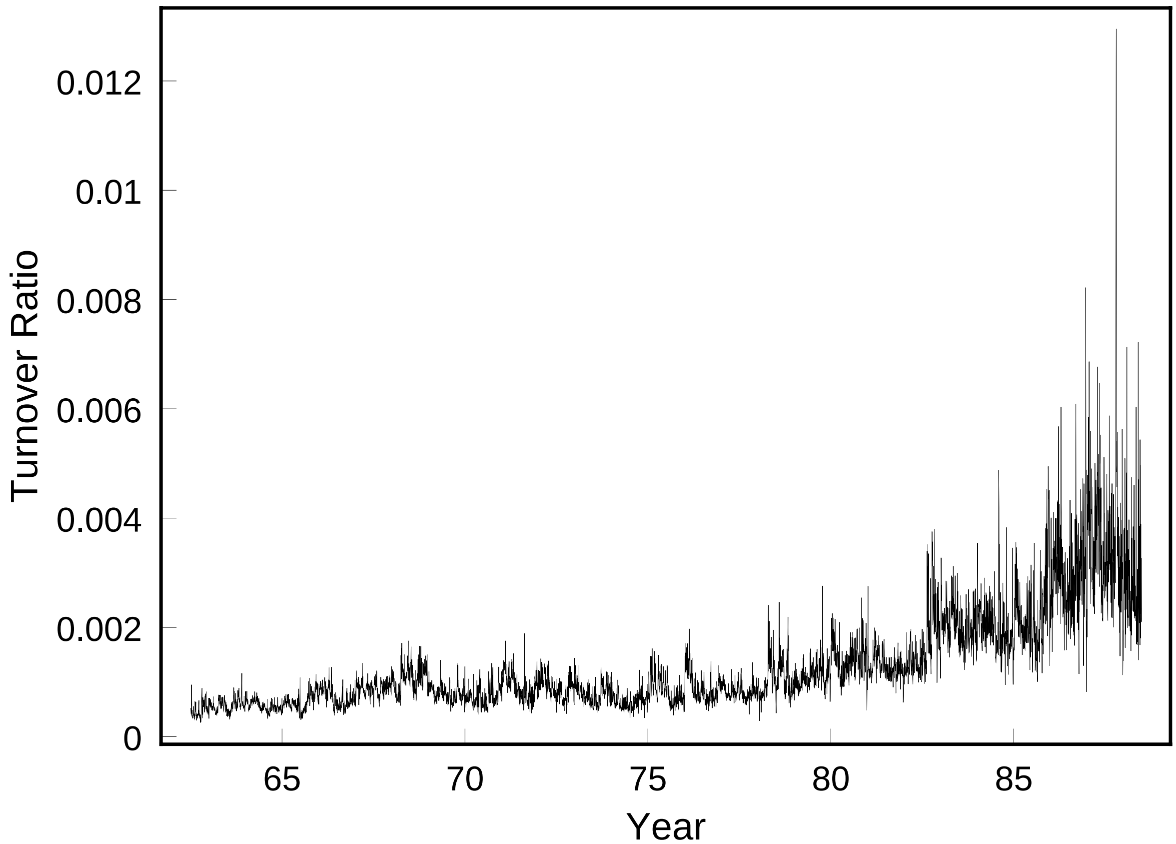

Figure 1: Daily NYSE Turnover Ratios July 62-June 88

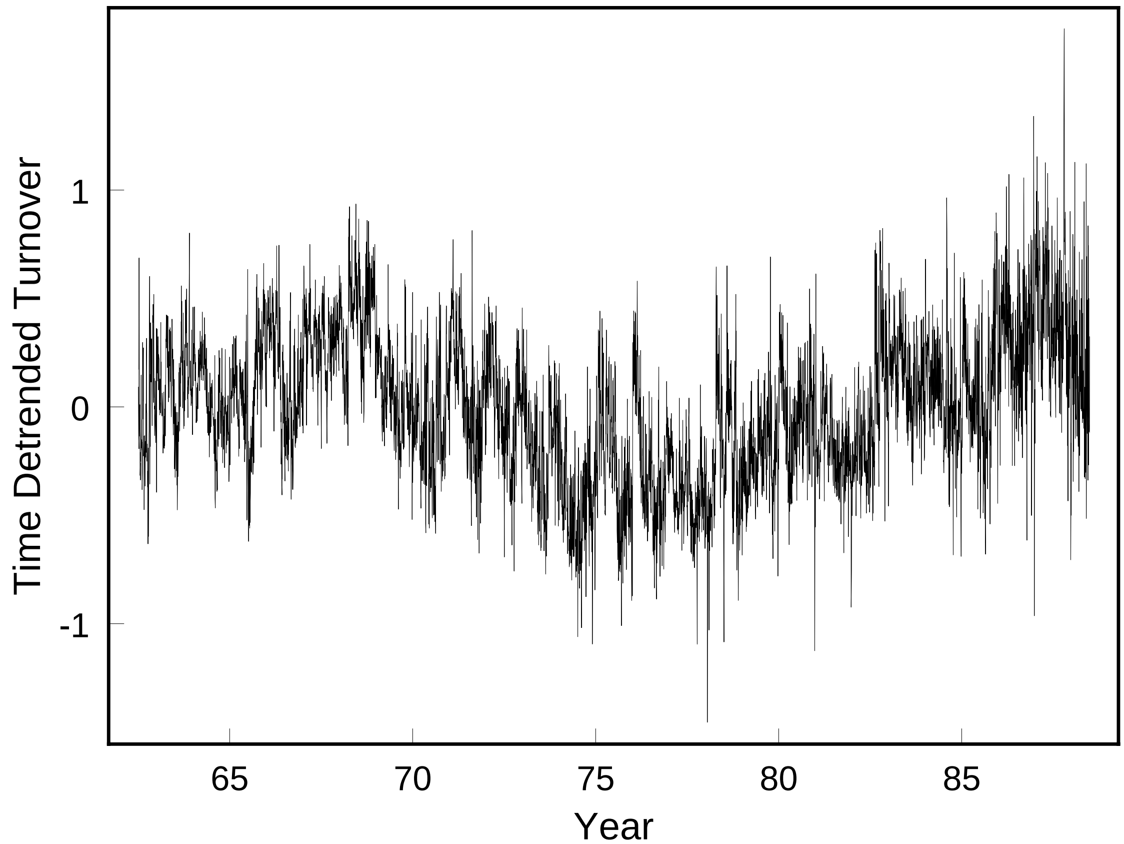

Figure 2: Linear Time Detrended Log(Turnover)

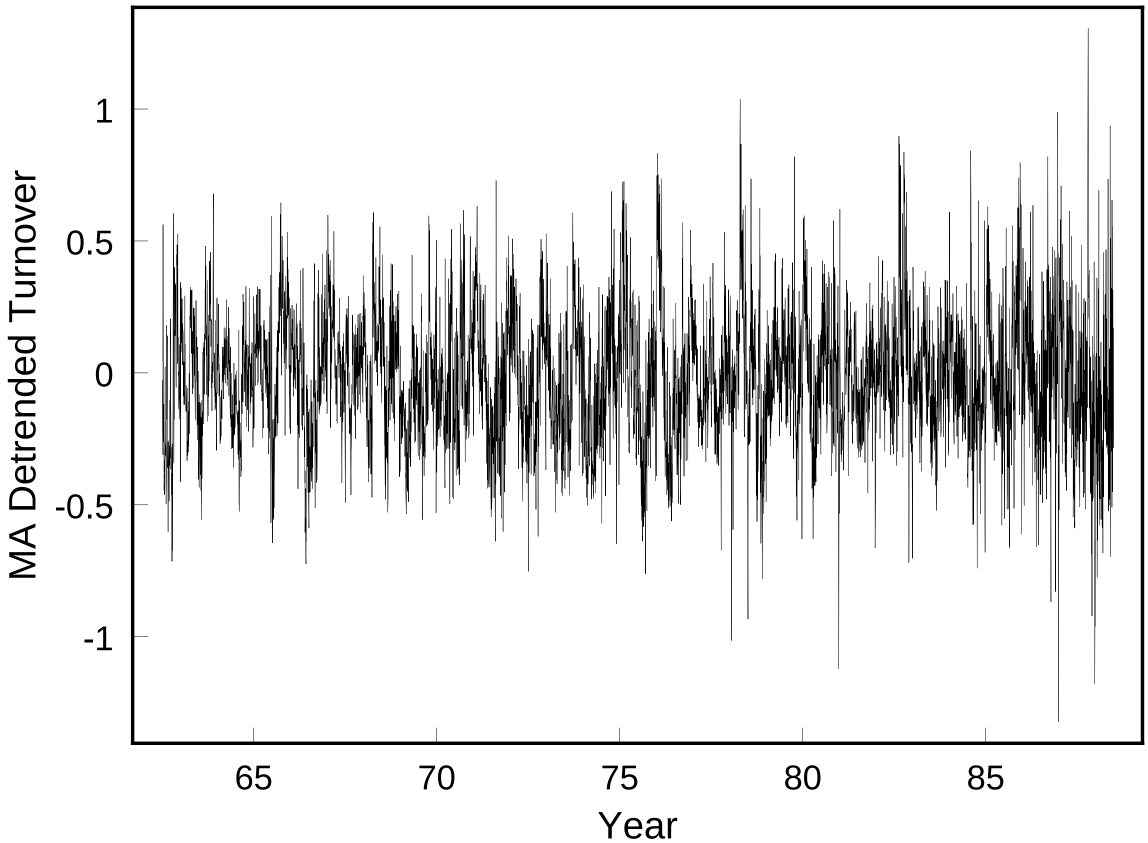

Figure 3: Log 100 Day Moving Average Detrended Turnover [Log(V/MA)]

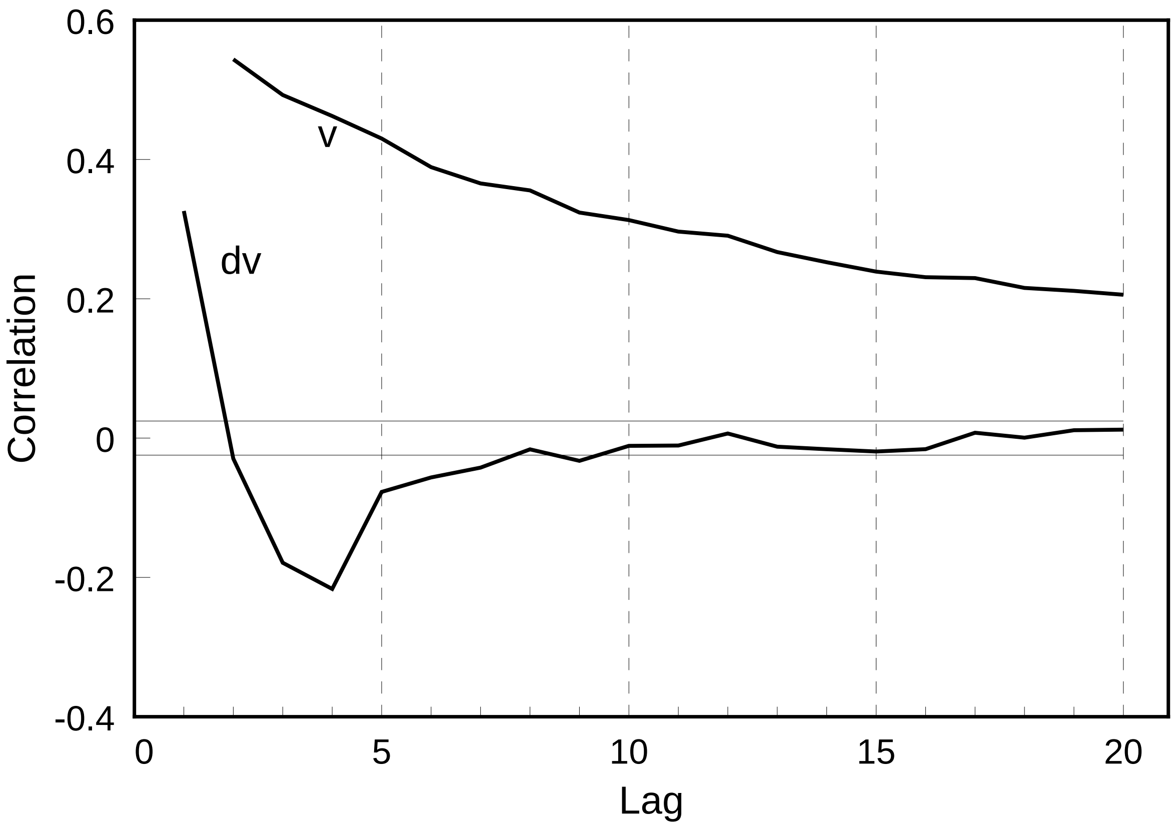

Figure 4: Volume Autocorrelations. v = volume normalized using 100 day ma, dv = volume normalized using 5 day ma, and r = index return.

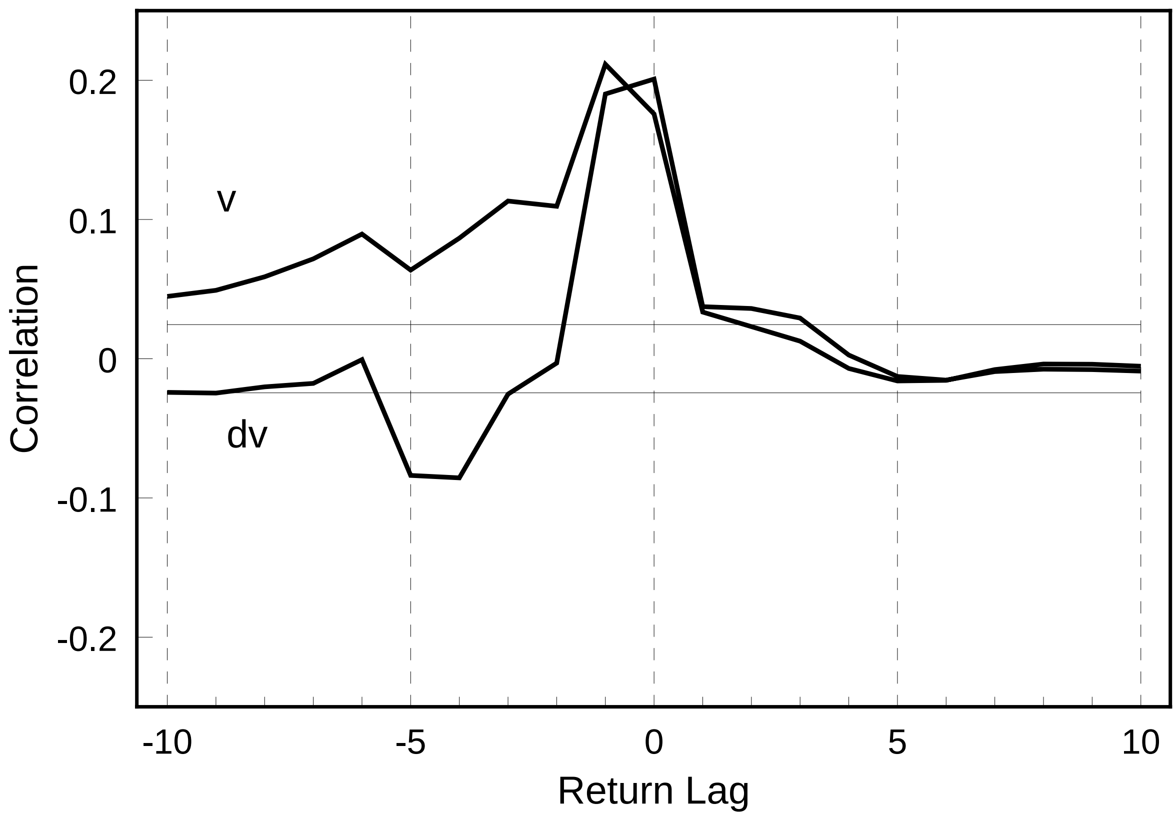

Figure 5: Return Volume Cross Correlations. Correlations from v(t), and dv(t) with r(t+j). v = volume normalized using 100 day ma, dv = volume normalized using 5 day ma, and r = index return.

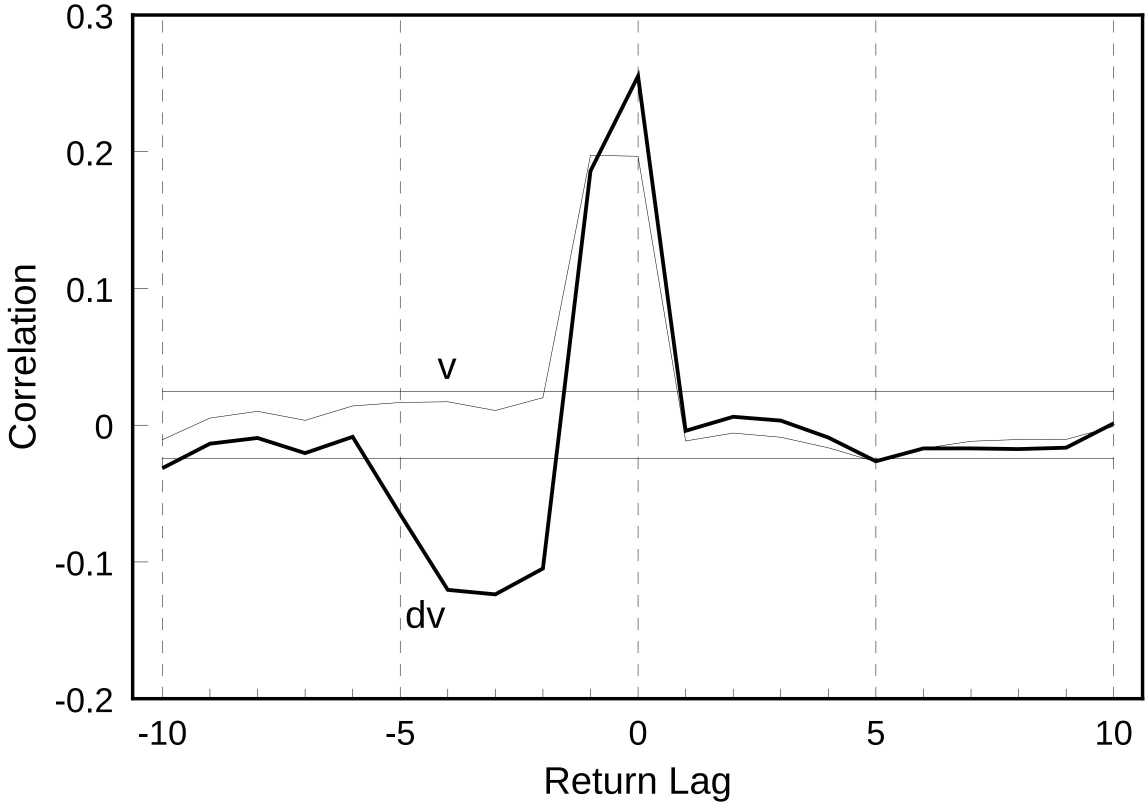

Figure 5a: Return Volume Cross Correlations. Correlations from v(t), and dv(t) with |r(t+j)|. v = volume normalized using 100 day ma, dv = volume normalized using 5 day ma, and r = index return.

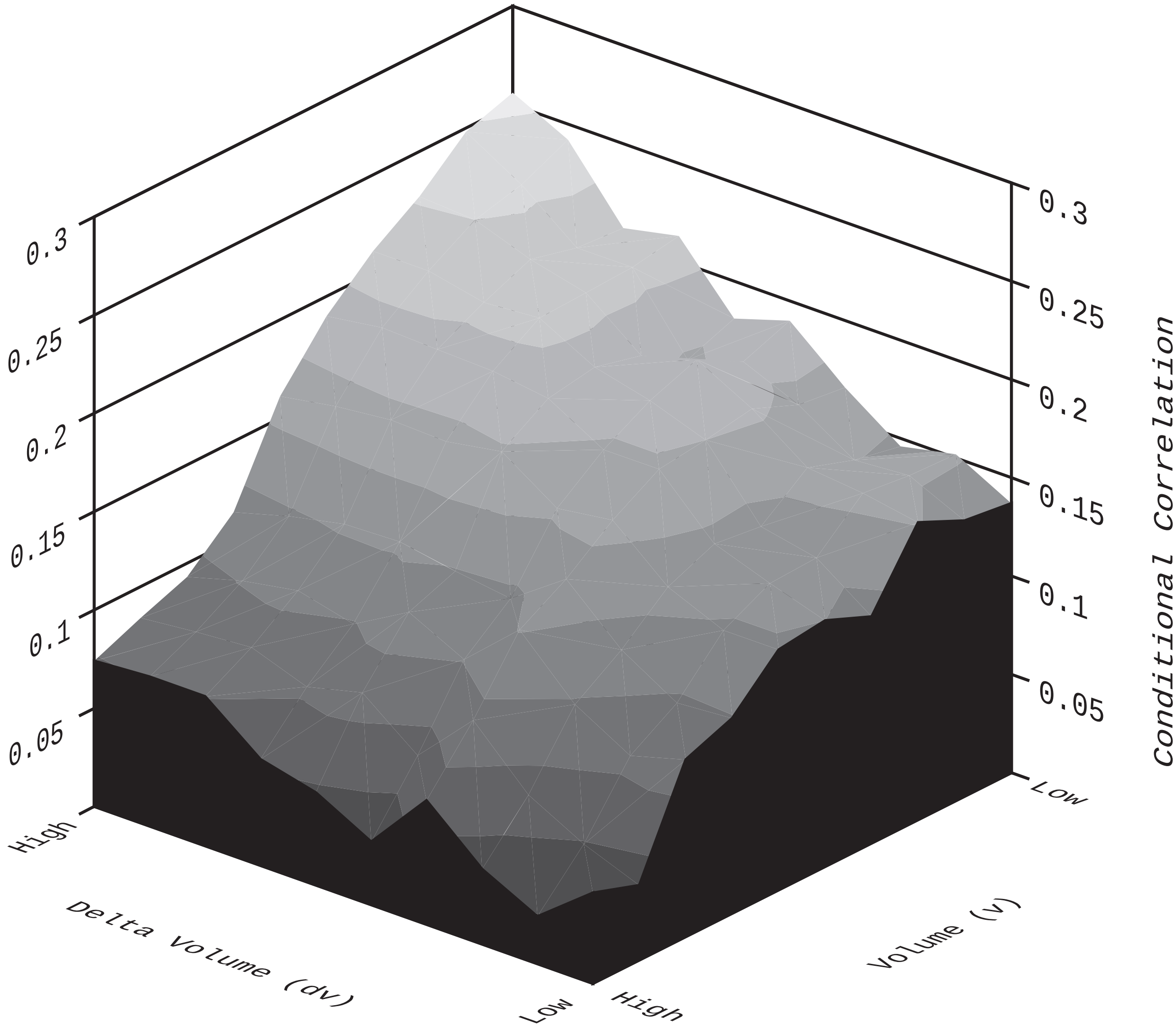

Figure 6: Estimated conditional correlation r(t), r(t+1) using v(t) and dv(t) as conditioning information. Estimation is done using a uniform kernel of bandwidth 0.3 on fractile transformed v and dv series.

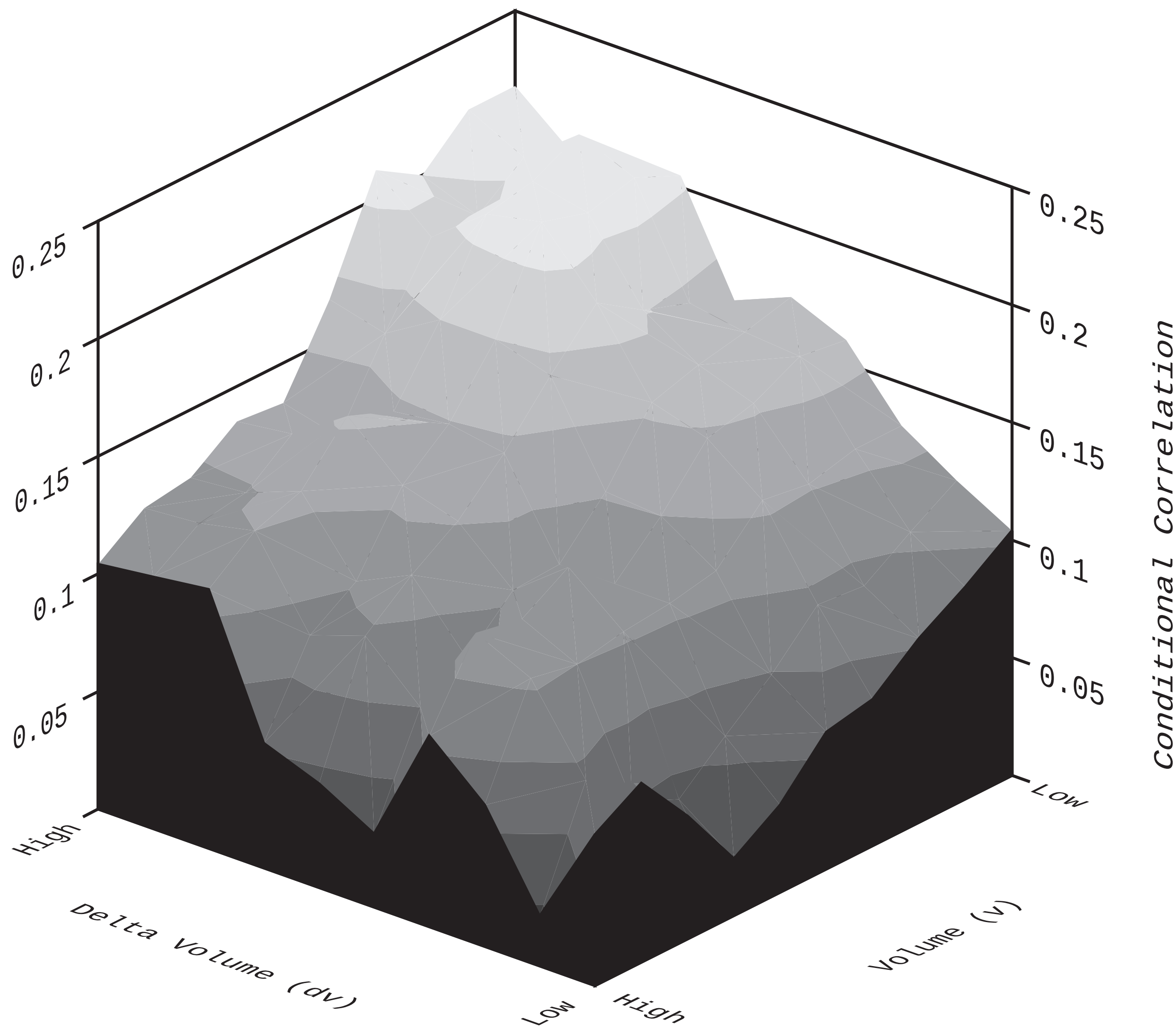

Figure 7: Estimated conditional correlation r(t), r(t+1) using v(t) and dv(t) as conditioning information restricting r(t)>0. Estimation is done using a uniform kernel of bandwidth 0.3 on fractile transformed v and dv series.

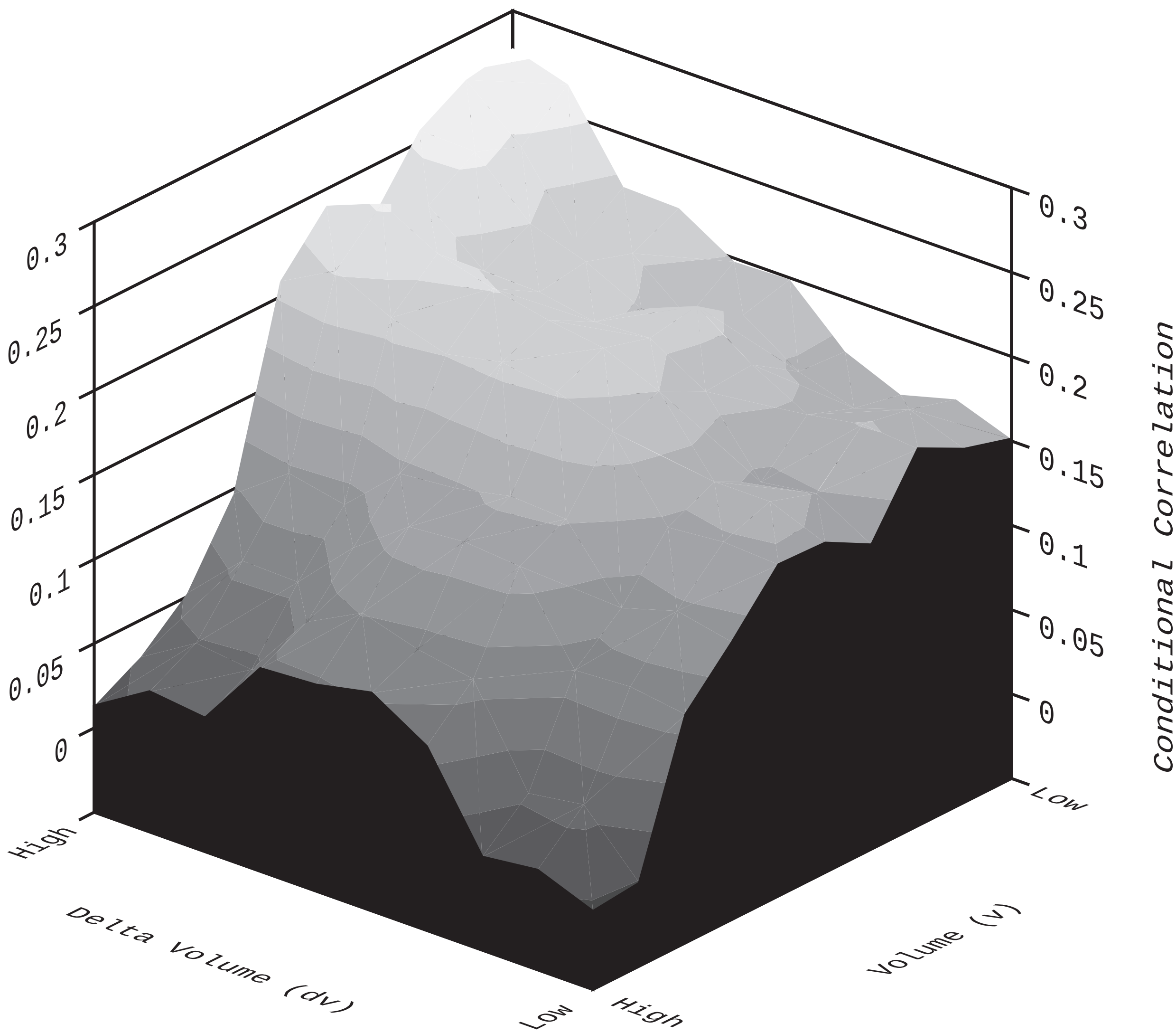

Figure 8: Estimated conditional correlation r(t), r(t+1) for r(t)<0, using v(t) and dv(t) as conditioning information. Estimation is done using a uniform kernel of bandwidth 0.3 on fractile transformed v and dv series.

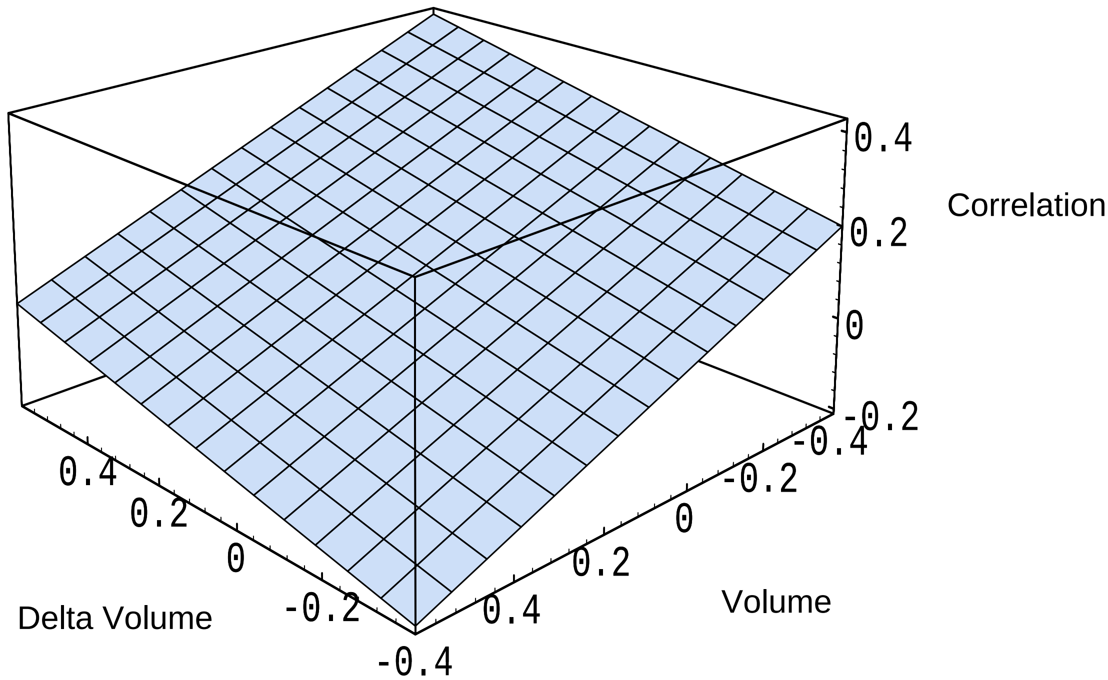

Figure 9: Fitted conditional correlation (positive and negative returns). Coefficients are from table 4. Volume ranges are central 98% of the distributions.

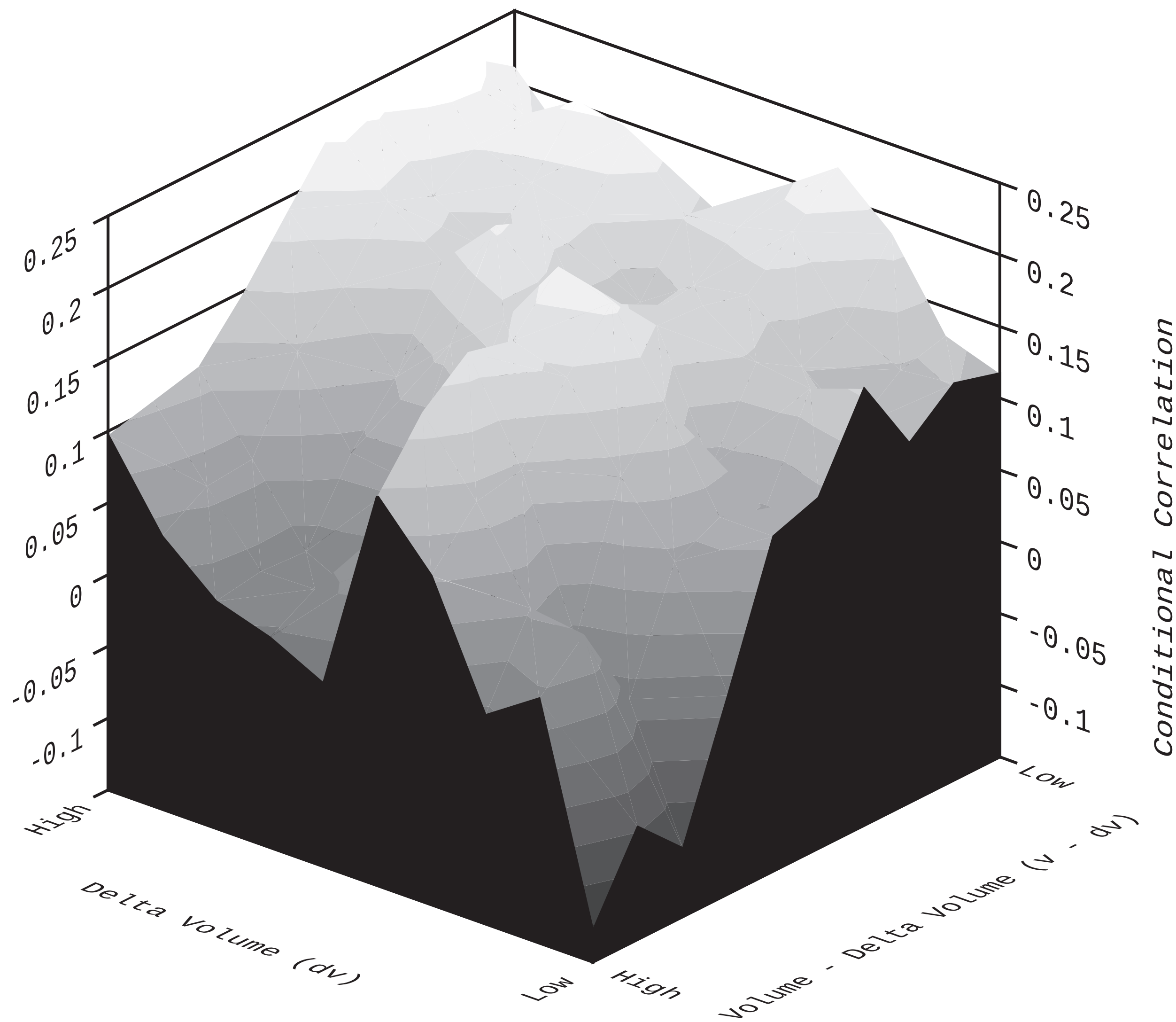

Figure 10: Estimated conditional correlation r(t), r(t+1) for r(t)<0, using v(t) - dv(t) and dv(t) as conditioning information. Estimation is done using a uniform kernel of bandwidth 0.3 on fractile transformed v and dv series.

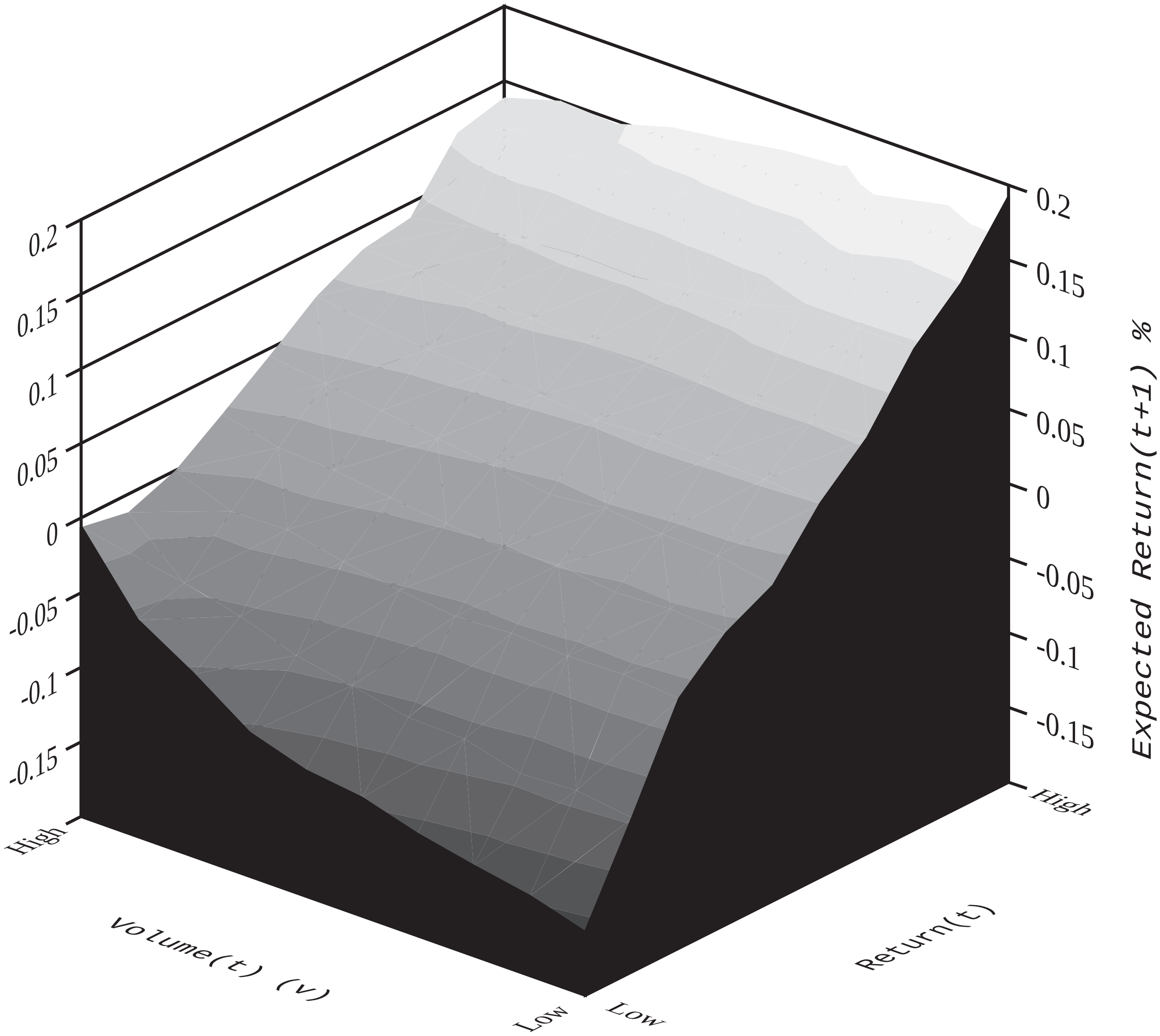

Figure 11: Estimated expected return r(t+t), using v(t) and r(t) as conditioning information. Estimation is done using a uniform kernel with bandwidth equal to 1.3 and 1.2 standard deviations for volume and returns, respectively. Bandwidth was determined using cross-validation squared error minimization.

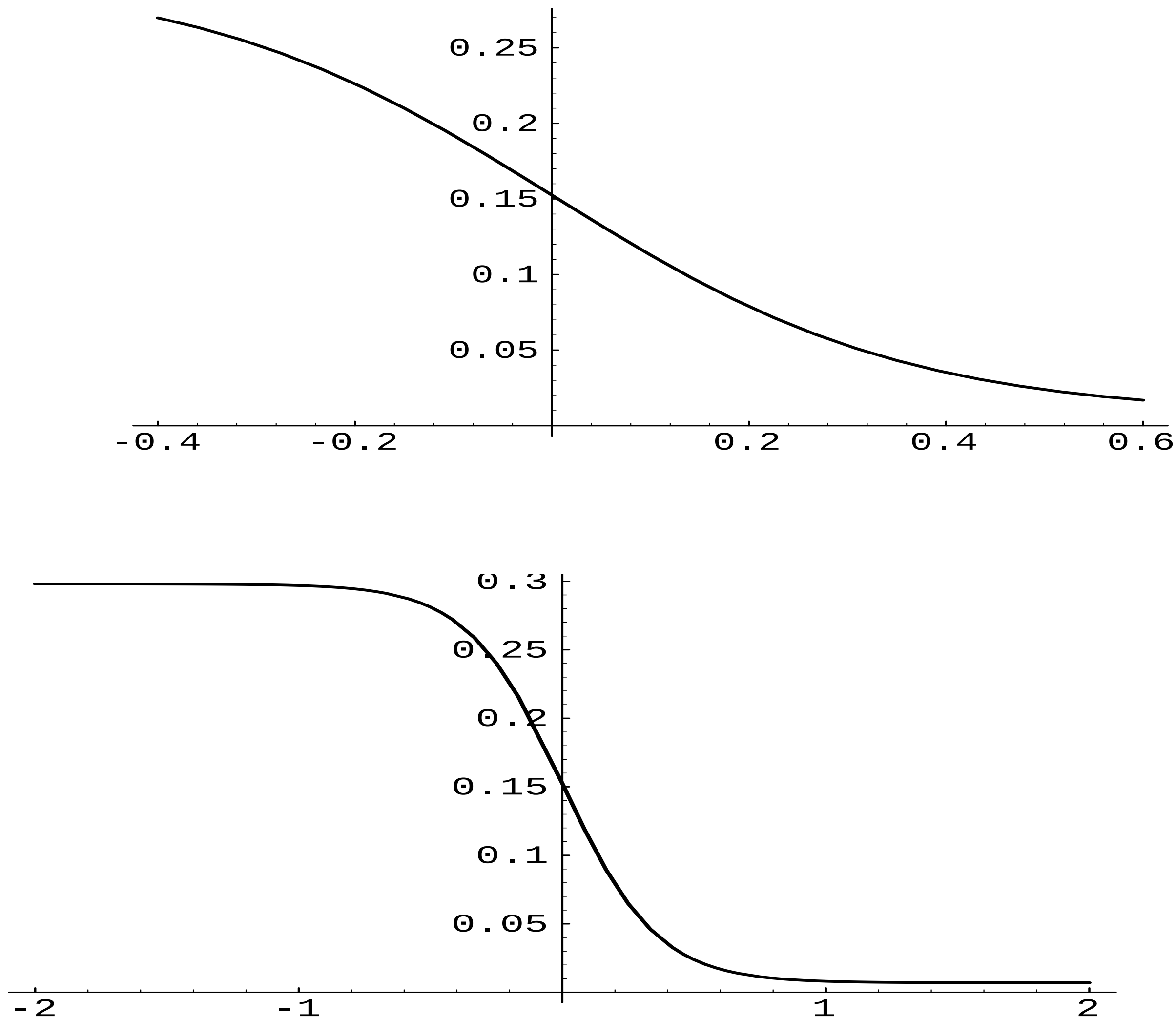

Figure 12 (Panels A and B): Estimated smooth threshold function over central 98% of volume data (top panel), and over larger region (bottom panel).

Table 1
Firms: Sorted By Percentage Spread

Percentage Spread	Name	Ticker
0.001194	INTERNATIONAL BUSINESS MACHS	IBM
0.001810	GENERAL ELEC CO	GE
0.001843	INTERNATIONAL PAPER CO	IP
0.001907	MERCK & CO INC	MRK
0.002829	MINNESOTA MNG & MFG CO	MMM
0.002985	PROCTER & GAMBLE CO	PG
0.003328	AMERICAN TEL & TELEG CO	T
0.003431	GENERAL MTRS CORP	GM
0.003899	INCO LTD	N
0.003968	TEXACO INC	TX
0.004228	EXXON CORP	XON
0.004415	WOOLWORTH F W CO	Z
0.005231	CHEVRON CORPORATION	CHV
0.005348	UNITED TECHNOLOGIES CORP	UTX
0.005495	DU PONT	DD
0.005602	ALUMINUM CO AMER	AA
0.005634	WESTINGHOUSE ELEC CORP	WX
0.005666	AMERICAN BRANDS INC	AMB
0.005900	EASTMAN KODAK CO	EK
0.005917	GOODYEAR TIRE & RUBR CO	GT
0.006270	SEARS ROEBUCK & CO	S

Spread data is $2(a-b)/(b+a)$ for each firm sampled at 9:45 ET on September 9th, 1991.

Table 2
Daily Regressions

$$x_t = a + b_1 I_t(Mon) + b_2 I_t(Tue) + b_3 I_t(Thr) + b_4 I_t(Fri)$$

Series	a	Mon	Tue	Thu	Fri	R^2
Raw v	0.034	-0.120	-0.030	-0.012	-0.068	0.033
	(0.006)	(0.009)	(0.009)	(0.009)	(0.009)	
Raw dv	0.036	-0.121	-0.030	-0.010	-0.066	0.083
	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)	

Estimates for daily dummies. Estimation is by OLS and standard errors are OLS standard errors. The standard errors should viewed with some caution since the residuals are highly correlated. Residuals of this regression are used for all tests on the comovements of volume and returns.

Table 3
Summary Statistics

Series	Dow21	VW	v	dv
Mean*100	0.040	0.047	0.000	0.000
Std*100	0.867	0.780	23.210	14.614
Skewness	0.254	0.132	0.222	-0.290
Kurtosis	5.186	5.730	3.771	7.278
Max	0.054	0.053	1.122	0.791
Min	-0.051	-0.044	-1.289	-1.593
ρ_1	0.127	0.218	0.691	0.326
ρ_2	-0.004	0.017	0.544	-0.031
ρ_3	-0.007	0.025	0.493	-0.179
ρ_4	-0.018	0.002	0.462	-0.219
ρ_5	-0.016	0.001	0.430	-0.078
Bartlett Std.	0.0125	0.0125	0.0125	0.0125
Corr(Dow Index)	0.966	0.945		

Summary statistics for the 21 firm index (Dow21), and the Value Weighted CRSP Index(VW). Both series include dividends. The two volume series (v and dv) are detrended using a 100, and 5 day MA respectively. They also have had day of the week effects removed. (See table 2). ρ_i is the autocorrelation at lag i. Corr(Dow Index) is the contemporaneous correlation with the Dow Jones Index.

Table 4
VW Comparisons

$$r_{t+1} = \alpha + (\beta_0 + \beta_1 v_t + \beta_2 S_t dv_t + \beta_3 \bar{S}_t dv_t) r_t$$
$$S_t = 1 \quad r_t \geq 0, \quad S_t = 0 \quad r_t < 0$$
$$\bar{S}_t = 1 \quad r_t < 0, \quad \bar{S}_t = 0 \quad r_t \geq 0$$

Series	β_0	β_1	β_2	β_3	R^2
Dow21 62-87					
AR(1)	0.127				0.016
	(0.016)				
v	0.153	-0.268			0.021
	(0.015)	(0.069)			
dv	0.139	-0.396	0.426		0.023
	(0.017)	(0.091)	(0.143)		
dv neg	0.138	-0.402	0.433	0.028	0.023
	(0.017)	(0.093)	(0.142)	(0.195)	
VW 62-87 (Sept)					
AR(1)	0.218				0.048
	(0.016)				
v	0.250	-0.301			0.054
	(0.016)	(0.065)			
dv	0.241	-0.415	0.340		0.056
	(0.016)	(0.088)	(0.132)		
dv neg	0.242	-0.396	0.316	-0.090	0.056
	(0.016)	(0.087)	(0.131)	(0.193)	

Estimation is by OLS. Numbers in parenthesis are White heteroskedasticity consistent standard errors. Dow21 is the 21 Dow firm index. VW is the CRSP value weighted index.

Table 5
Subsamples and Specification Tests

$$r_{t+1} = \alpha + (\beta_0 + \beta_1 v_t + \beta_2 S_t dv_t) r_t$$
$$S_t = 1 \quad r_t \geq 0, \quad S_t = 0 \quad r_t < 0$$

Series	β_0	β_1	β_2	R^2
Dow21 62-74	0.214	-0.613	0.622	0.054
	(0.026)	(0.164)	(0.234)	
Dow21 75-87 (Sept)	0.078	-0.257	0.308	0.009
	(0.021)	(0.113)	(0.178)	
Dow21: Volume Rank Transform	0.136	-0.329	0.225	0.023
	(0.017)	(0.067)	(0.077)	
Kernel Residuals	0.049	-0.287	0.350	0.006
	(0.017)	(0.093)	(0.147)	
$dv = AR(5)$ Residuals	0.141	-0.394	0.306	0.023
	(0.017)	(0.096)	(0.127)	
$dv = v$ Residuals	0.143	-0.285	0.382	0.022
	(0.016)	(0.071)	(0.157)	

Rank tranform tranforms the two volume series to their fractile rankings. Kernel residuals uses the estimates the above expression on the residuals of a kernel estimated of the expected return conditioned on lagged return and lagged v. dv = AR(5) Residual sets the dv series the the residual of an AR(5) fit to v. $dv = v$ residuals sets dv equal to residuals of the raw series regressed on contemporaneous v. Estimation is by OLS. Numbers in parenthesis are White heteroskedasticity consistent standard errors.

Table 6
Persistance at Longer Lags

$$r_{t+j} = \alpha + (\beta_0 + \beta_1 v_t + \beta_2 S_t dv_t) r_t$$
$$S_t = 1 \quad r_t \geq 0, \quad S_t = 0 \quad r_t < 0$$

Series	β_0	β_1	β_2	R^2
$j = 2$	-0.003	-0.029		0.000
	(0.016)	(0.062)		
	-0.016	-0.127	0.390	0.002
	(0.017)	(0.074)	(0.139)	
$j = 3$	-0.012	0.061		0.000
	(0.016)	(0.066)		
	-0.019	0.007	0.179	0.001
	(0.017)	(0.072)	(0.145)	

Estimation is by OLS. Numbers in parenthesis are White heteroskedasticity consistent standard errors.

Table 7
Smooth Threshold Estimation

$$r_{t+1} = \alpha + (\beta_0 + \frac{\beta_1}{1 + e^{\beta_2 v_t}} + S_t(\beta_3 + \frac{\beta_4}{1 + e^{\beta_5 d v_t}}))r_t$$

$$S_t = 1 \quad r_t \geq 0, \quad S_t = 0 \quad r_t < 0$$

Series	β_0	β_1	β_2	β_3	β_4	β_5	R^2
Dow21	0.007	0.291	5.58				0.021
	(0.085)	(0.172)	(5.05)				
Dow21	-0.106	0.505	3.88	0.147	-0.335	7.49	0.024
	(0.194)	(0.385)	(3.69)	(0.133)	(0.267)	(8.45)	

Estimation is by NLLS. Numbers in parenthesis are White heteroskedasticity consistent standard errors.

Table 8
Conditional Probabilities $r_t r_{t+1} > 0$

$r_t > 0$

	$v_t < 0$		$v_t > 0$
$dv_t < 0$	0.557	(0.50)	0.550
	(3.60)	(0.48)	(0.15)
	[1013]		[462]
$dv_t > 0$	0.605	(4.16)	0.552
	[537]		[1371]

$r_t < 0$

	$v_t < 0$		$v_t > 0$
$dv_t < 0$	0.529	(3.00)	0.490
	(4.59)	(0.90)	(2.07)
	[1288]		[520]
$dv_t > 0$	0.589	(4.98)	0.519
	[511]		[831]

Probabilities of positive or negative returns continuing from t to t+1. Numbers in parenthesis are t-statistics testing equality of fraction with appropriate neighbor in the table. Numbers in brackets are the number of observations in each cell.

Table 9
Individual and Market Residuals

$$r_{t+1} = \alpha + (\beta_0 + \beta_1 v_t + \beta_2 S_t dv_t) r_t$$
$$S_t = 1 \quad r_t \geq 0, \quad S_t = 0 \quad r_t < 0$$

Series	β_0	β_1	β_2	R^2
Individual Residuals	0.072	-0.272	0.239	0.007
	(0.017)	(0.091)	(0.146)	
Market Residuals	0.107	0.008	-0.130	0.011
	(0.015)	(0.076)	(0.178)	

Model fitted to residuals of nonlinear model fit to each individual firm, and mean of market model residuals for each firm using 1000 day rolling beta. Estimation is by OLS. Numbers in parenthesis are White heteroskedasticity consistent standard errors.

Table 10
Portfolio Leads

$$r_{i,t+1} = \alpha + (\beta_0 + \beta_1 v_t + \beta_2 S_t dv_t) r_{j,t}$$
$$S_t = 1 \quad r_t \geq 0, \quad S_t = 0 \quad r_t < 0$$

Series	β_0	β_1	β_2	R^2
Low Beta on Dow21	0.149	-0.278	0.129	0.024
	(0.016)	(0.091)	(0.139)	
Medium Beta on Dow21	0.141	-0.396	0.616	0.022
	(0.020)	(0.105)	(0.177)	
High Beta on Dow21	0.140	-0.432	0.420	0.014
	(0.022)	(0.115)	(0.181)	
Small on Small	0.128	0.025	0.067	0.008
	(0.022)	(0.112)	(0.196)	
Small on Large	0.042	-0.201	0.318	0.014
	(0.008)	(0.042)	(0.073)	

First portfolio regressed on lag of second. Beta estimates are rolling over previous 1000 days. Small and large refer to small and large move portfolios. On a market rise large refers to the half of the Dow21 group with the largest return. On a market fall it refers to the group with the smallest return. Estimation is by OLS. Numbers in parenthesis are White heteroskedasticity consistent standard errors.

Table 11
Conditional Expected Returns(t+1) for Various Information Sets

Rule	N Returns	Trades	Mean$*10^3$	Std.$*10^3$	t
All	6345		0.40	8.67	
$r_t > 0$	3383	2965	1.31	9.03	
$r_t > 0.004$	2010	2519	1.62	9.51	1.18
$r_t > 0.004, v_t < 0$	828	1291	1.65	9.83	0.91
$r_t > 0.004, v_t < 0, dv_t > 0$	339	564	2.16	9.71	1.55

Conditional returns and standard deviations using lagged return and volume information. Trades is the number of trades (both in and out of the market) that would be undertaken by a hypothetical trading strategy. t is a t-statistic test for equality of the given conditional mean and the mean for $r_t > 0$.