Persistent Intraday Correlations Create Skews in Daily-Scale Distributions Revisited: Behaviors of High-Magnitude Fluctuations

T. J. Mazurek

t_j_mazurek@yahoo.com

May 3, 2018

Abstract. This work began with an earlier study that showed the extreme persistence of intraday correlations in order-flow and price-mobility creates positively skewed daily-scale distributions of bull-bear sentiment and price-change, respectively. The earlier work also dealt with the all-magnitude skew excesses as one of each distribution's summary parameters and compared scaling behaviors of all corresponding parameters of price-change and sentiment. The present study extends this prior work by determining the price-mobility and order-flow persistent correlations for high magnitude price-changes. This study shows that the bias in daily occurrence probabilities at high-magnitudes determined from price-mobility correlations favors negative price-changes, reversing the positive bias of the all-magnitude probabilities. This produces a negative daily price-mobility skew that creates an equal negative daily price-change distribution's skew excess at high magnitudes. The behavior of high-magnitude order-flow correlations parallels the behavior of price-mobility described above to create a similar negative sentiment distribution skew at high-magnitudes. The present study then compares behaviors with increasing fluctuation magnitudes of price-change and sentiment skews on the daily and seven-day scale. These scaling comparisons show that the daily negative skew strength at high-magnitudes is amplified at successively higher scales as was the daily positive all-magnitude skew strength in the earlier work. Such amplifications sharply contrast with behaviors of skew strengths for independent identical distributions which must decay with increases in scale.

1. Introduction

This work began with an earlier study (cf., Mazurek 2017[1]) which showed that extremely persistent intraday sign-correlations of order-flow[2] and price-mobility[3] create positive skews in the distributions of bull-bear sentiment[4] and price-change in the following manner. Order-flow auto correlations characterize the tendency of buy (sell) orders to be followed more buy (sell) orders. Price-mobility auto correlations characterize the tendency of positive (negative) price-changes to be followed by more positive (negative) price-changes. Both correlations are extremely persistent in the sense that their auto correlations are positive and statistically significant out to lags of thousands of trades or price-changes, corresponding to trading times that span over several days. These respective persistent correlations were then used to determine the daily occurrence probabilities of buyer and seller trades and positive and negative price-changes. The daily occurrence probabilities determined from order-flow were biased in favor of buyer trades, while the ones determined from price-mobility were biased toward positive price-changes. These respective biases then created positive skews in both the daily sentiment and the daily price-change distributions. The earlier study also introduced the normal inverse Gaussian representation of empirical distributions which completely specifies a distribution by its summary parameters and used this representation to examine the scaling behaviors with increasing scales of the sentiment and price-change distributions. It found that for both bull-bear sentiment and price-change the daily distributions' skews were amplified as the scales increased.

The present study extends the earlier work by deriving the intraday price-mobility and order-flow auto correlations for high-magnitude price-changes. The high-magnitude (Hi-Mag) regimes for sentiment and price-change are defined as the regions where the negative density distributions are greater than the positive ones. The specific Hi-Mag thresholds are determined in Section 3 and are different for the distributions of sentiment and price-change.

The Hi-Mag individual sign-correlations persist at statically significant levels out to lag times of a couple days, so their persistency is somewhat lower than the all magnitude[5] (All-Mag) correlations which remain statistically significant out to lags of several days. The corresponding average of positive and negative price-changes for price-mobility, and the average of buyer and seller correlations for order-flow have lower persistency levels as did the corresponding All-Mag results. The average of individual correlations for price-mobility remain statistically significant

---

[1] The currently posted version of this paper corrects errors in an earlier version originally posted in July 2017. The author become aware of some remaining transcription errors in the paper as now posted and will email a corrected version on request.

[2] The extreme persistence of order-flow correlations was first pointed out independently by Bouchaud, Gefen, Potters, and Wyart (2004), and Lillo and Farmer (2004). Bouchaud, Farmer, and Lillo (2009) present a review of such long memory processes.

[3] The earlier study introduced price-mobility and showed that it is as extremely persistent as is order-flow.

[4] The daily bull-bear sentiment is the difference of the buyer minus seller daily trade totals.

[5] The all magnitude correlations include all of the fluctuations.

out to a lags of about a day, while the corresponding average for order-flow is statistically significant out to lags of roughly half of a day.

First consider the Hi-Mag price-mobility results. The Hi-Mag price-mobility auto correlations behave contrarily to the All-Mag ones examined in the earlier study: here the probability of a negative price-change to be followed by more negative price-changes dominates the corresponding sequence for positive, the reverse of the All-Mag case. This work determines the individual probabilities for occurrence of Hi-Mag positive and negative price-changes from the individual positive and negative Hi-Mag price-mobility correlations. These probabilities then define the daily Hi-Mag price-mobility skew. The results show this Hi-Mag daily price-mobility skew is biased in favor of negative price-changes, the opposite of the positive price-change bias for the All-Mag case.

The Hi-Mag order-flow behavior parallels that of price-mobility as described in the prior paragraph. This work determines the individual probabilities for occurrence of a Hi-Mag buy and sell trades from the individual Hi-Mag order-flow auto correlations. Again the Hi-Mag order-flow auto correlations behave contrarily to the All-Mag ones examined in the earlier study: the probability of a sell trade to be followed by more sell trades dominates the corresponding sequence for buy trades and the daily occurrence probability of a seller trade is greater than that of a buyer one. Hence the daily order-flow skew excess is biased toward sellers, the opposite of the buyer trade bias for the All-Mag case.

Let's now look at the implications of both reversed biases on the daily scale Hi-Mag distribution skews for price-change and sentiment. The above contrary behaviors of Hi-Mag skews to the All-Mag ones parallel the same contrary behaviors of daily-scale price-change and sentiment distribution skew excesses: both Hi-Mag skew excesses are negative while the All-Mag ones are positive. The probability bias toward negative price-changes described in the second of above paragraphs creates daily Hi-Mag price-changes that are also biased toward negative prices-changes. One thus expects that the daily Hi-Mag price-change distribution will have a negative skew. Similarly the above paragraph's bias toward seller trades creates a bias for Hi-Mag daily trade totals to be greater for sellers than those of buyers. This creates negative Hi-Mag daily skews in the sentiment distribution since sentiment is the difference of buyer minus seller trade-totals. This explains why the empirical Hi-Mag skews of sentiment and price-change distributions exhibit exactly both of these reversals in signs from their All-Mag daily distribution skews.

Let's now compare the Hi-Mag price-mobility and order-flow daily skew excesses to their respective price-change and sentiment distributions' skew excesses. Section 2 determines the daily Hi-Mag price-mobility skew excess to be $\varepsilon_{daily}^{PM,HiMag} = -0.066 \pm 0.002$. The daily Hi-Mag price-change skew excess as given in Section 3 is $\varepsilon^{\Delta lnP^c}(1.2) = -0.105 \pm 0.039$. Hence to within one standard deviation $\varepsilon_{daily}^{PM,HiMag}$ equals $\varepsilon_1^{\Delta lnP^c}(1.2)$ and the Hi-Mag price-change fluctuations carry the imprints of the Hi-Mag intraday price-mobility correlations. Section 2 also gives the daily Hi-Mag order-flow skew excess $\varepsilon_{daily}^{OF,HiMag} = -0.035 \pm 0.002$. The Hi-Mag skew excess for sentiment given in Section 3 is $\varepsilon_1^{N_B-N_S}(0.8) = -0.078 \pm 0.061$. The latter two quantities' one-standard deviation regions overlap. Hence the daily Hi-Mag sentiment fluctuations carry the imprint of the Hi-Mag intraday auto correlations.

This work then goes on to present and compare the daily positive and negative branches of the cumulative distributions of sentiment and price-change and finds that the two distributions are nearly identical. Next it examines the daily scale behaviors of the price-change and sentiment skew excesses with increasing threshold magnitudes. Both skew excesses decrease roughly linearly with increasing threshold magnitudes in essentially identical manners. The correlation of the behaviors with increasing magnitudes of the sentiment and price-change excesses is 88%.

This work next considers the seven-day scale positive and negative branches of the cumulative distributions of sentiment and price-change. The price-change cumulative distributions show marked increases in skews with obvious widenings of the separations between the positive and negative branches over those on the daily scale. This indicates appreciable growth of the seven-day scale skews above their initial daily strengths. The sentiment cumulative distributions do not openly exhibit increases in skew due to classification errors as noted at the end of the next paragraph.

Finally this work compares the daily and seven-day scale price-change and sentiment skew excesses behaviors with increasing threshold magnitudes. For price-change both the positive and negative skew excesses at the seven-day scale show significant amplification above their daily scale values. The positive skews increase by factors greater than one and up to two above their daily scale values, while the negative ones increase by factors greater than one and up to three. Thus the initial daily skews for price-change amplify significantly as the scales increase. A similar amplification with increasing scales also occurred for the All-Mag price-change skew in the earlier study. On the other hand comparisons of the Hi-Mag skew excesses for sentiment to their daily values in the present study do not show amplification of skews. For positive skews on the seven-day scale, there are more skews with values below their daily ones than there are with values above their daily ones. The negative skews actually show a trend of skew-magnitude shrinkage from daily values: the negative seven-day skews systematically lie above the negative daily ones. As discussed in more detail in the earlier study, the anemic growth there of the all-magnitude skews and the lack of growth and shrinkage here for higher magnitude fluctuations in the seven-day sentiment skews is due to errors introduced by classification method used to determine buyer and seller trades.

The remainder of this paper is structured as follows. Section 2 presents the Hi-Mag intraday sign-correlations results for price-mobility and order-flow for various lag times. This section also determines the daily Hi-Mag price-mobility skew excess and the daily Hi-Mag order-flow skew excess. Section 3 defines the positive and negative branches of the cumulative distributions above given threshold magnitudes for price-change and bull-bear sentiment. It uses these cumulative distributions to define skew excess for higher magnitude fluctuations above given threshold magnitudes. This section then discusses the behaviors of the two branches of the cumulative distributions as well as the skew excess' behavior with increasing threshold magnitudes for both sentiment and price-change on the daily and seven day scales. Section 4 summarizes the results and conjectures how trader activities give rise to the Hi-Mag negative biases in price-mobility and order-flow that produce the negative skews in price-change and sentiment. The Appendix describes the sign auto correlations produced by skewed independent distributions.

2. Long-Term Intraday Sign-Correlations for Price-Mobility and Order-Flow at High-Magnitudes

This section defines and presents the high-magnitude (Hi-Mag) price-mobility and order-flow long-term intraday sign-correlations. It also determines the Hi-Mag daily price-mobility and order-flow skew excesses.

To determine each of these Hi-Mag auto correlations one must first remove the overwhelmingly dominant correlations produced by the fluctuations below the high magnitude demarcation boundaries. These boundaries are determined in Section 3, and are different for price-change and sentiment. This study removed the dominant correlations by replacing each respective empirical low-magnitude fluctuation by a corresponding symmetric and independent-sign fluctuation having unity magnitude.[6] Because the signs of the latter fluctuations are symmetric and uncorrelated their net contribution to auto correlations is zero, leaving only the auto correlation due to the Hi-Mag empirical fluctuations.

2.1 Long-Term Intraday Sign-Correlations for Hi-Mag Price-Mobility

The Hi-Mag lower boundary is determined in Section 3 to be given by $1.2\sigma_1^{\Delta lnP^c}$ where $\sigma_1^{\Delta lnP^c} = 0.0106$ is the daily price-change volatility. To determine the Hi-Mag auto correlation one keeps the original empirical trades only when the intraday price P_{i_t} satisfies the minimum price-change condition

[6] In C++, suitable sets of symmetric and independent-sign fluctuations can be generated from the random number generator rand() by dividing each of its outputs by the maximum number it generates ($RAND\_MAX$), and assigning a fluctuation of -1 for each result less than one half and $+1$ for each otherwise result.

$$\left| lnP_{i_t} - lnP_{i-1}^c - \langle ln(P_j^c/P_{j-1}^c)\rangle \right| \geq 1.2\sigma_1^{\Delta lnP^c}, \tag{2.1.1}$$

where $\langle ln(P_j^c/P_{j-1}^c)\rangle = 0.000466$ is the mean or drift and P_{i-1}^c is the prior day's closing price. Each fluctuation that fails to satisfy equation (2.1.1) is replaced by an independent-sign symmetric fluctuation as described in footnote 6. The procedures for determining the price-mobility correlations for this transformed data set at lag time τ are detailed in Subsection 2.7 of the earlier study (Mazurek 2017). This now determines the Hi-Mag auto correlations since the substituted independent-sign fluctuations give no net contribution.

Figure 2.1.1 below gives the individual price-mobility sign-correlations $C_{++}^{HiMag}(\tau)$, $C_{--}^{HiMag}(\tau)$, and the average price-mobility sign-correlation $\bar{C}_{PM}^{HiMag}(\tau) = -C_{+-}^{HiMag}(\tau)$. Labels appear on some correlations in the figure designating nominal[7] lagged clock times starting at about 13 minutes and moving progressively higher out to two days.

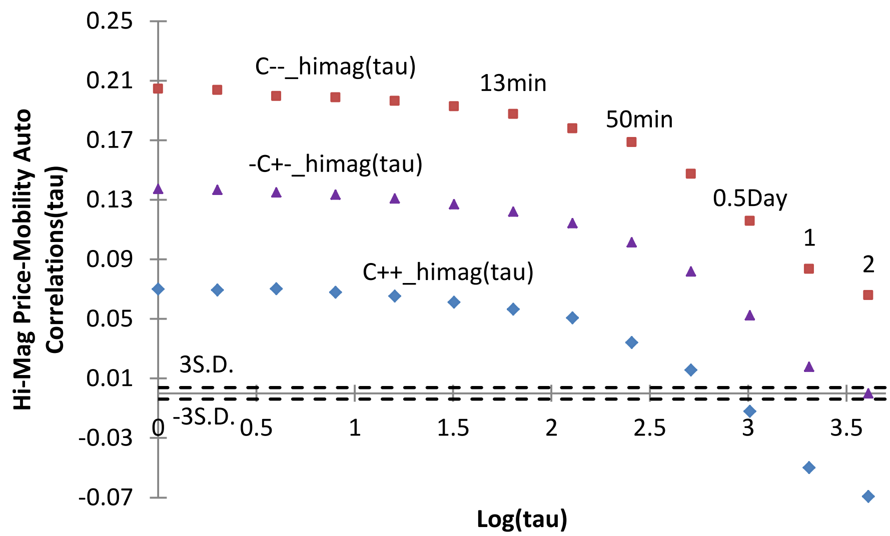

Figure 2.1.1 Individual Hi-Mag price-mobility sign-correlations $C_{++}^{HiMag}(\tau)$, $C_{--}^{HiMag}(\tau)$, and the average price-mobility sign-correlations $\bar{C}_{PM}^{HiMag}(\tau) = -C_{+-}^{HiMag}(\tau)$.

The three standard deviation (± 0.0039) dotted lines run parallel to the horizontal axis and correspond to the individual price-mobility correlations[8]. The behaviors of the Hi-Mag price-mobility correlations have roughly similar general behaviors with increasing lags as those of the All-Mag ones in the earlier study, although their starting values are much smaller and ending values at large τ are appreciably different. The most striking difference from the All-Mag auto

[7] The clock times correspond to averages and are nominal in the sense that their corresponding standard deviations may be comparable to the averages or even larger.

[8] The average price-mobility correlations have a lower standard deviation of ± 0.0019.

correlations are that the dominant correlations are reversed here so that $C_{--}^{HiMag}(\tau) > C_{++}^{HiMag}(\tau)$ at all lags τ.

The individual correlations of $C_{++}^{HiMag}(\tau)$ and $C_{--}^{HiMag}(\tau)$ at the highest lag have reached roughly equal magnitudes but have differing signs with the respective values of about -0.069 and 0.066. The corresponding All-Mag correlations also had equal constant magnitudes with different signs at highest lags. The Appendix shows that this type of behavior with equal correlation magnitudes but opposite signs for the individual sign correlations arises from skewed independent fluctuation distributions.

Let's now consider the probabilities of positive and negative price-changes at lags corresponding to the daily scale for $\tau \sim 10^{3.3} \sim 2000$. Using equations (2.7.10) of Subsection 2.7 in the earlier work (Mazurek 2017) one obtains the daily positive and negative price-change probabilities of $P_{+}^{HiMag} = 0.467 \pm 0.001$ and $P_{-}^{HiMag} = 0.533 \pm 0.001$, respectively, where for simplicity the lag designation is suppressed. The daily Hi-Mag price-mobility skew excess is then given by

$$\varepsilon_{daily}^{PM,HiMag} = P_{+}^{HiMag} - P_{-}^{HiMag} = -0.066 \pm 0.002 \qquad (2.1.2)$$

Let's compare the above daily Hi-Mag price-mobility with the Hi-Mag price-change skew excess determined in Section 3 for the longest duration IVE data set $\varepsilon_1^{\Delta lnP^c}(1.2) = -0.105 \pm 0.039$. One sees that the one-standard deviations regions overlap so that the two skews are effectively equal. Thus the Hi-Mag fluctuations carry the imprint of the intraday price-mobility correlations.

2.2 Long-Term Intraday Sign-Correlations for Hi-Mag Order-Flow

The Hi-Mag lower boundary for sentiment is determined in Section 3 to be $0.8\sigma_1^{N^B-N^S}$, where the daily sentiment volatility is $\sigma_1^{N^B-N^S} = 300.2$. To determine the Hi-Mag auto correlation one keeps the original empirical trades only when the intraday totals $N_{i_t}^B - N_{i_t}^S$ satisfy the minimum price-change condition

$$\left| N_{i_t}^B - N_{i_t}^S - \langle N_j^{N^B-N^S} \rangle_1 \right| \geq 0.8\sigma_1^{N^B-N^S}, \qquad (2.2.1)$$

where $\langle N_j^{N^B-N^S} \rangle_1 = 49.8$ is the daily sentiment mean or drift. Each fluctuation that fails to satisfy equation (2.2.1) is replaced by an independent-sign fluctuation as described in footnote 6. The procedures for determining the price-mobility correlations for this transformed data set at lag time τ are detailed in Subsection 2.7 of the earlier study (Mazurek 2017). This now determines the Hi-Mag auto correlations for sentiment since the substituted independent-sign fluctuations give no net contribution.

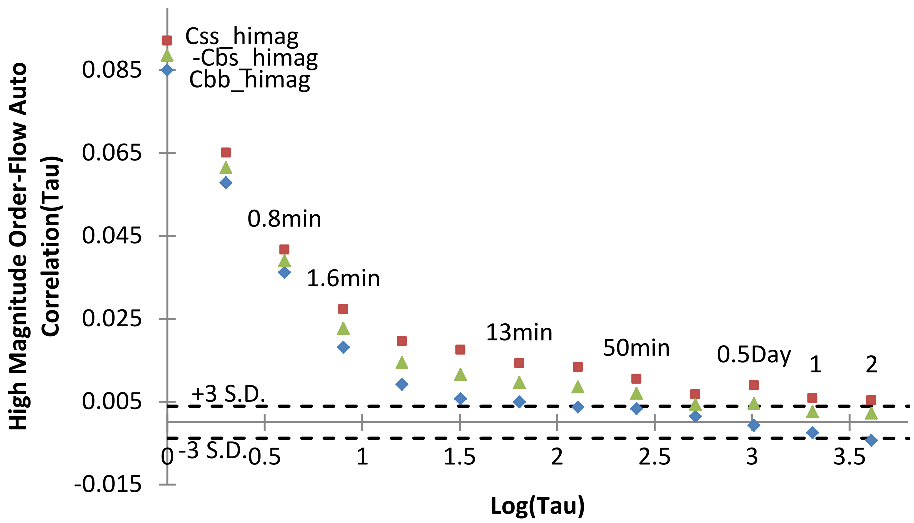

Figure 2.2.1 Individual Hi-Mag order-flow sign-correlations $C_{BB}^{HiMag}(\tau)$, $C_{SS}^{HiMag}(\tau)$, and the average order-flow sign-correlations $\bar{C}_{OF}^{HiMag}(\tau) = -C_{BS}^{HiMag}(\tau)$.

Figure 2.2.1 above gives the individual order-flow sign-correlations $C_{BB}^{HiMag}(\tau)$, $C_{SS}^{HiMag}(\tau)$, and the average price-mobility sign-correlation $\bar{C}_{OF}^{HiMag}(\tau) = -C_{BS}^{HiMag}(\tau)$.[9] Labels appear on some correlations in the figure designating nominal[10] lagged clock times starting at about 0.8 minutes and moving progressively higher out two days. The three standard deviation (± 0.0039) dotted lines are parallel to the horizontal axis and correspond to the individual order-flow correlations[11]. The behaviors of the Hi-Mag order-flow correlations have roughly similar behaviors with increasing lags as those of the All-Mag ones in the earlier study, although their starting values are much smaller and ending values at the largest lags are different. As in the case of price-mobility above, The most striking difference from the All-Mag auto correlations are that the dominant correlations are reversed here so that $C_{SS}^{HiMag}(\tau) > C_{BB}^{HiMag}(\tau)$ at all lags τ.

As was the case for price-mobility, the individual correlations of $C_{BB}^{HiMag}(\tau)$ and $C_{SS}^{HiMag}(\tau)$ at the highest lag reached roughly equal magnitudes with differing signs and have the respective values of -0.004 and 0.005. The corresponding All-Mag correlations in the earlier study also had constant magnitudes with opposite signs at the highest lags. Such correlation behavior again

[9] All these correlations are Lee-Ready quantities (cf. Subsection 2.4 in Mazurek 2017) but for notational simplicity the LR superscript is suppressed here.

[10] The clock times correspond to averages and are nominal in the sense that their corresponding standard deviations may be comparable to the averages or even larger.

[11] The average order-flow correlations have a lower standard deviation of ± 0.0019.

indicates that the highest lag correlations are coming from skewed independent order-flow fluctuations (cf. Appendix).

Let's now consider the probabilities of buyer and seller trades at lags corresponding to the daily scale for $\tau \sim 10^{3.3} \sim 2000$. Using equation (2.4.8) of Subsection 2.4 in the earlier work (Mazurek 2017) one obtains the daily buyer and seller trade probabilities $P_B^{HiMag} = 0.483 \pm 0.002$ and $P_S^{HiMag} = 0.518 \pm 0.002$, respectively, where for simplicity the lag designation is suppressed. One then obtains the following daily skew for order-flow.

$$\varepsilon_{daily}^{OF,HiMag} = P_B - P_S = -0.035 \pm 0.004 \qquad (2.2.2)$$

Now compare this daily Hi-Mag order-flow skew excess with the Hi-Mag price-change skew excess determined in Section 3 for the IVE data set $\varepsilon_1^{N_B - N_S}(0.8) = -0.078 \pm 0.061$. One sees that the one-standard deviation regions overlap. Thus the high magnitude price-change fluctuations carry the imprint of the Hi-Mag order-flow correlations.

3. Positive and Negative Branches of Cumulative Distributions and Skew Excess Behaviors with Increasing Threshold Magnitudes

This section defines the positive and negative branches of the cumulative distributions of bull-bear sentiment and price-change. It uses these to define the skew excess above a given threshold magnitudes and to set the lower boundary of the high magnitude fluctuation regions for both distributions. This section also discusses the behaviors of the two branches of the cumulative distributions as well as the skew excess' behavior with increasing threshold magnitudes for both sentiment and price-change on the daily and seven day scales.

To consider in more detail the empirical skew characteristic of the logarithmic price-changes, one needs to define suitable skew parameters. To prepare for this definition let's first consider the positive and negative branches of the cumulative distribution for price-change. For a given scale s, the these branches of the cumulative distribution at or above a given magnitude m_s are given by

$$F_{\pm,s}^{\Delta lnP^c}\left(m_s/\sigma_s^{\Delta lnP^c}\right) \;=\; \frac{1}{N_s^{tot}} \sum_{i=1}^{N_s^{tot}} \Theta\left(\pm z_{i,s}^{\Delta lnP^c}/\sigma_s^{\Delta lnP^c} - m_s/\sigma_s^{\Delta lnP^c}\right), \qquad (3.1)$$

where $z_s^{\Delta lnP^c}$ is the drift-free price-change[12], $\sigma_s^{\Delta lnP^c}$ the volatility, N_s^{tot} is the number of price-changes at scale s, and $\Theta(y)$ is the unit step function that equals zero for $y < 0$ and unity

[12] The drift-free price-changes on the daily and higher scales are detailed in Subsections 2.5 and 2.6 of the earlier study (Mazurek 2017).

otherwise. Now define the price-change skew excess above magnitude m_s using the cumulative distributions at or above this given threshold magnitude as

$$\varepsilon_S^{\Delta lnP^c}\left(m_s/\sigma_S^{\Delta lnP^c}\right) = \frac{F_{+,S}^{\Delta lnP^C}\left(m_s/\sigma_S^{\Delta lnP^C}\right) - F_{-,S}^{\Delta lnP^C}\left(m_s/\sigma_S^{\Delta lnP^C}\right)}{F_{+,S}^{\Delta lnP^C}\left(m_s/\sigma_S^{\Delta lnP^C}\right) + F_{-,S}^{\Delta lnP^C}\left(m_s/\sigma_S^{\Delta lnP^C}\right)}\ . \qquad (3.2)$$

The corresponding functions for the bull-bear sentiment variable are obtained via the substitutions: $\Delta lnP^c \rightarrow N^B - N^S$ in equations (3.1) and (3.2) where N^B (N^S) denote buyer (seller) trade totals.

The lower boundary of the high magnitude region can be set by the following considerations. For continuous distributions, the cumulative distributions are expressed in terms of their probability density functions $f_\pm(m/\sigma)$ as $F_\pm(m/\sigma) = \int_{m/\sigma}^{\infty} f_\pm(y)\, dy$. For simplicity the latter equations suppress the designations for sentiment, price-change, and scale. The lower boundary of the Hi-Mag regions is then defined as the point where the difference $F_+(m/\sigma) - F_-(m/\sigma)$ takes on its minimum value[13]. At this point $m_{\min[F_+ - F_-]}$, the positive and negative density functions are equal and $-f_+\left(m_{\min[F_+ - F_-]}/\sigma\right) + f_-\left(m_{\min[F_+ - F_-]}/\sigma\right) > 0$ for $m > m_{\min[F_+ - F_-]}$. This minimum point is the lower boundary of the region where negative density distribution dominates the positive one for the continuous case. The daily high magnitude skew excess is now defined as $\varepsilon\left(m_{\min[F_+ - F_-]}/\sigma\right)$. The latter skew characterizes the Hi-Mag fluctuations in the same way as all of the fluctuations are characterized by the All-Mag skew $\varepsilon(0)$. These considerations should also apply to the empirical distributions and this study uses the minimums of the sentiment and price-change cumulative distribution differences to define the Hi-Mag boundary.

Using closing price data[14] of the iShares S&P 500 Value ETF with ticker symbol IVE that has a total duration of $N_{tot} = 4347$ days or about 17.4 years for price-change one obtains $m_{\min[F_+ - F_-]}^{\Delta lnP^c} = 1.2\sigma_1^{\Delta lnP^c}$ and the daily price-change skew excess for the Hi-Mag region is $\varepsilon_1^{\Delta lnP^c}(1.2) = -0.105 \pm 0.039$.

The determination of sentiment Hi-Mag demarcation point uses the IVE data set described in Subsection 2.1 of the earlier study (Mazurek 2017) which has a shorter duration of about five years to obtain $m_{\min[F_+ - F_-]}^{N^B - N^S} = 0.8\sigma_1^{N^B - N^S}$ and the daily sentiment skew excess for the Hi-Mag region is $\varepsilon_1^{N_B - N_S}(0.8) = -0.078 \pm 0.061$. All following results in this section also use the latter data set.

[13] This procedure fails for symmetric distributions where the difference of the cumulative distributions vanishes everywhere, leaving the Hi-Mag boundary undefined.

[14] The closing prices are from Yahoo! Finance and the data set spans the duration from 5/26/2000 to 9/8/2017.

Let's now look at the behaviors with increasing threshold magnitudes of the sentiment and price-change cumulative distributions and skew excesses. Figure 3.1 below shows the daily scale positive and negative branches of the cumulative distributions of price-change and sentiment (the latter is plotted displaced downward by one unit). One sees that to within the resolution in this figure both distributions are essentially the same.

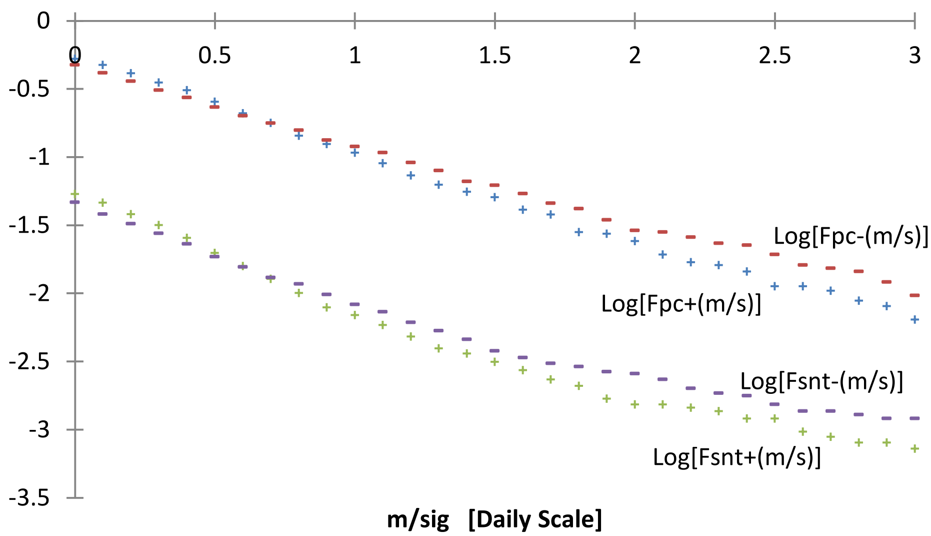

Figure 3.1 $F_{\pm,1}^{\Delta lnP^c}\left(m_1/\sigma_1^{\Delta lnP^c}\right) = \text{Fpc}\pm(\text{m/s})$ and $F_{\pm,1}^{N^B-N^S}\left(m_1/\sigma_1^{N^B-N^S}\right) = \text{Fsnt} \pm (\text{m/s})$, on the daily scale. The sentiment plot is displaced downward by one unit.

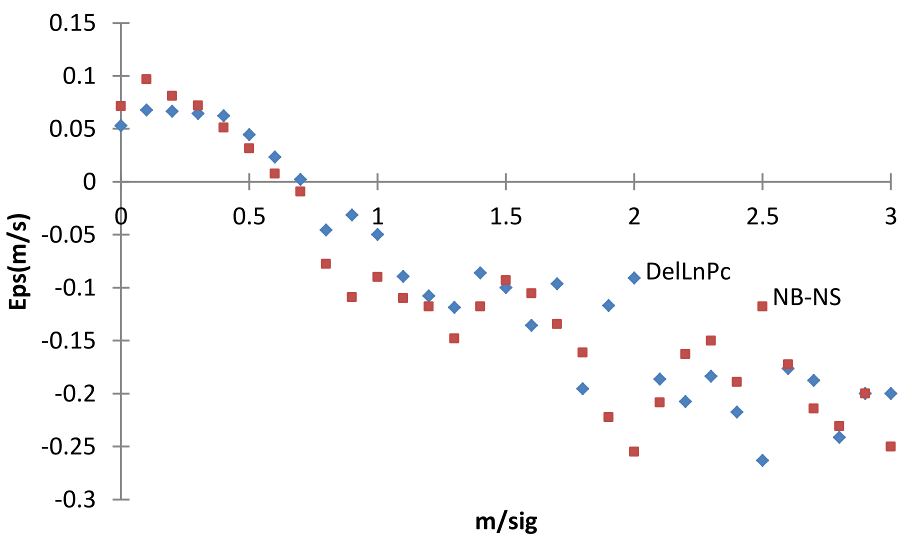

Figure 3.2 Variations of $\varepsilon_s^{\Delta lnP^c}\left(m_s/\sigma_s^{\Delta lnP^c}\right)$ and $\varepsilon_s^{N^B-N^S}\left(m_s/\sigma_s^{N^B-N^S}\right)$ on the daily scale.

Figure 3.2 gives variation of the sentiment and price-change skew excesses with increasing threshold magnitudes. There is considerable scatter in this Figure 3.2 due to the small sample-sizes. However the behaviors of the two skews decrease roughly linearly with increasing fluctuation magnitudes and are quite similar. Figure 3.3 below shows the variations of $\varepsilon_S^{N^B-N^S}(m_s/\sigma_S^{N^B-N^S})$ with $\varepsilon_S^{\Delta lnP^c}(m_s/\sigma_S^{\Delta lnP^c})$ including the trend line and its R^2 value. The correlation between the two skews is $R = 88\%$.

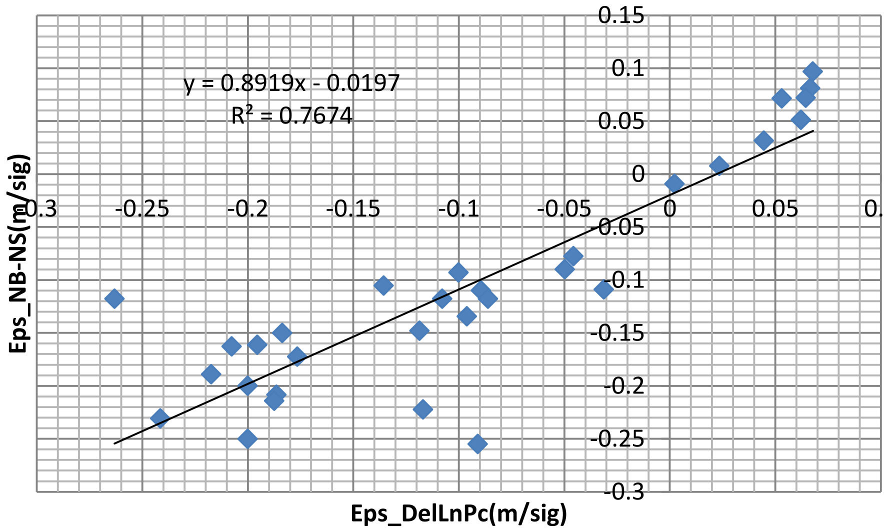

Figure 3.3 Correlation between $\varepsilon_S^{\Delta lnP^c}(m_s/\sigma_S^{\Delta lnP^c})$ and $\varepsilon_S^{N^B-N^S}(m_s/\sigma_S^{N^B-N^S})$ on the daily scale.

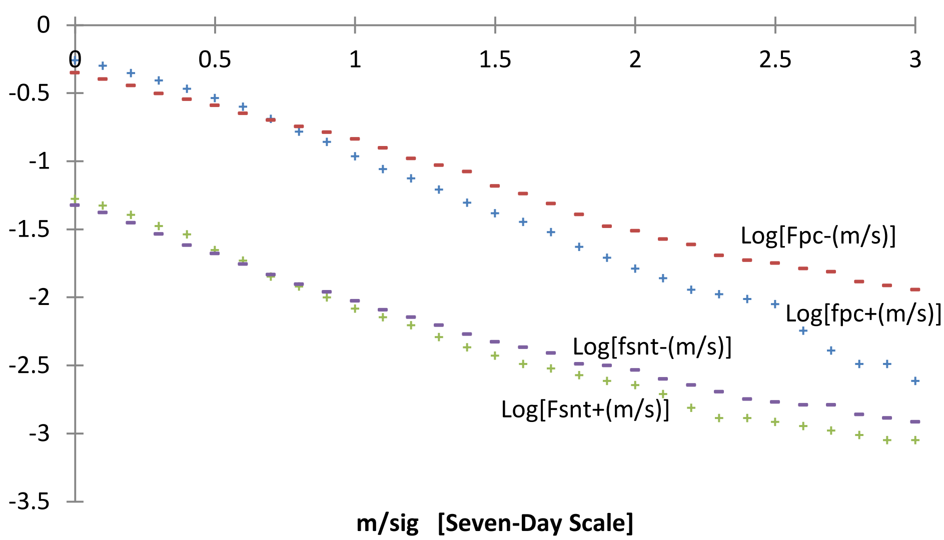

Figure 3.4 $F_{\pm,7}^{\Delta lnP^c}(m_7/\sigma_7^{\Delta lnP^c})$ = Fpc±(m/s) and $F_{\pm,7}^{N^B-N^S}\left(m_7/\sigma_7^{N^B-N^S}\right)$ = Fsnt ± (m/s) , on the seven-day scale. The sentiment plot is displaced downward by one unit.

Figure 3.4 shows the seven-day scale positive and negative branches of the cumulative distributions of price-change and sentiment (with the latter plotted displaced downward by one unit). One sees that the separations between the positive and negative branches of the price-change cumulative distributions has widened appreciably relative to daily separations everywhere, but especially at high magnitudes above $\sim 2.5\sigma_7^{\Delta lnP^c}$. Thus skews at all thresholds for price-change have markedly increased above the corresponding ones at the daily scale. On the seven-day scale the bull-bear sentiment cumulative distribution's positive and negative branches are essentially at the same separations as they were at the daily scale.

Figures 3.5 and 3.6 below compare the behaviors of the price-change and sentiment excesses with increasing fluctuation magnitudes on the daily and seven-day scales, respectively. The first figure shows that the price-change skew excesses at the seven-day scale amplify appreciably above their daily values everywhere they're non-zero. The amplification is greatest for negative skews at the highest magnitude thresholds. For positive skew excesses below $\sim 0.7\sigma_1^{\Delta lnP^c}$ the amplification factors over daily values are around two for most of the points; for negative skews above $\sim 0.9\sigma_1^{\Delta lnP^c}$ the amplification factors vary between 1.3 and 3.4.[15]

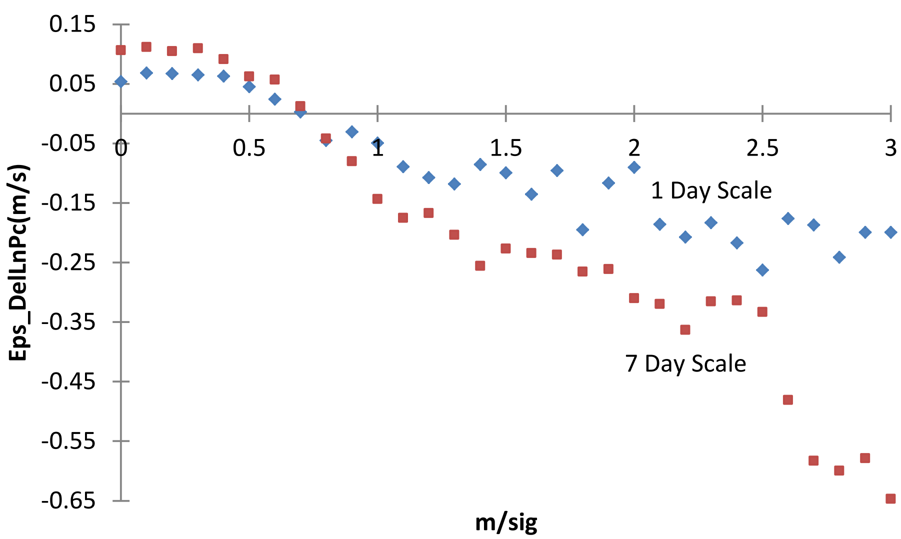

Figure 3.5 Comparing $\varepsilon_1^{\Delta lnP^c}\left(m_1/\sigma_1^{\Delta lnP^c}\right)$ and $\varepsilon_7^{\Delta lnP^c}\left(m_7/\sigma_7^{\Delta lnP^c}\right)$ with increasing thresholds.

A similar comparison of the seven-day and daily sentiment skews is shown in Figure 3.6 below. Within the relatively large scatter one can discern that the seven-day scale excesses systematically lie above those of the daily scale for thresholds above 0.8 times the volatility where they are negative. Thus the seven-day excess trend line lies above the daily one in these

[15] Using the S&P500 Index's much larger data set from 1/3/1950 to 2/3/2010 with closing prices from Yahoo! to derive price-change skew excessed similar to those in Figure 3.6 one obtains the following amplification factors: positive skews below $\sim 0.6\sigma_1^{\Delta lnP^c}$ have amplification factors that vary between 1.7 and 3, while negative skews above $\sim 0.7\sigma_1^{\Delta lnP^c}$ have amplification factors that vary between 2.3 and 7.3.

regions. This means that most of these excesses have magnitudes below the corresponding ones on the daily scale. The positive skews in the figure show varying amplification and diminishment with increasing scale, although the All-Mag skew does lie above the daily one at the seven day scale.

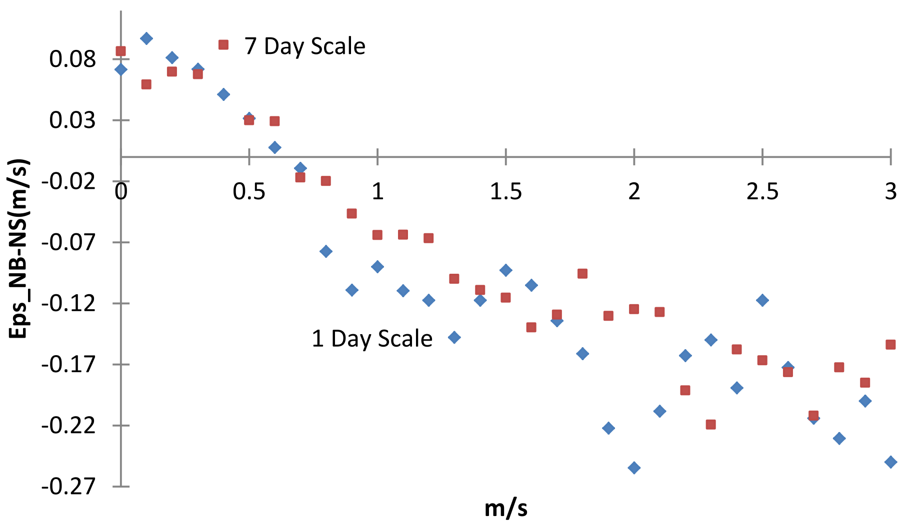

Figure 3.6 Comparing behaviors of $\varepsilon_1^{N^B-N^S}\left(m_1/\sigma_1^{N^B-N^S}\right)$ and $\varepsilon_7^{N^B-N^S}\left(m_7/\sigma_7^{N^B-N^S}\right)$ with increasing thresholds. .

The lack of amplification of the seven-day sentiment skews daily values is due to errors introduced by classification method used to determine buyer and seller trades. The Lee and Ready (1991) classification procedure was used here to identify buyer and seller trades. The earlier study (Mazurek 2017) introduces a true signal fidelity parameter $F_{TS>1.5LRN}(s)$ which represents the fraction of the series' components at each scale in which the true signal (TS) strongly dominates the Lee-Ready noise (LRN) . At the daily scale $F_{TS>1.5LRN}(s) = 0.7$. Thus at the daily scale only $\sim 70\%$ of the data set have strong signals relative to the noise. At the seven-day scale $F_{TS>1.5LRN}(s) = 0.5$, and only $\sim 50\%$ of the data have a strong signal. This accounts for the anemic rise over the daily scale value of the All-Mag sentiment skew and the decrease in strengths below daily values of the Hi-Mag sentiment skews with increasing scales.

4. Summary and Discussion

Section 2 presented the Hi-Mag intraday sign-correlations of price-mobility. Comparing the latter to the All-Mag price-mobility sign-correlations, one sees smaller auto correlations strengths for the former relative to the latter starting at one lag, followed by similar downward swings but to different ending values at the lag of two days. However both have extreme persistence out to a couple lag days. The striking difference between the two sets of auto correlations is that for the Hi-Mag order-flow correlations the probability for a negative price-change to be followed by more negative price-changes is stronger than probability of the corresponding sequence for positive price-changes. This reverses of the behaviors for the All-Mag case. Similarly for the Hi-Mag order-flow correlations the probability for a seller trade to be followed by another seller is greater than the corresponding sequence for buyer trades. This again is the reverse of the behaviors in the All-Mag case.

The existence of both these reversed auto correlations in the Hi-Mag data sets is remarkable. Recall that these data sets are created by replacing all intraday fluctuations below the high magnitude demarcation point by symmetric and independent-sign fluctuation which make no contribution to the auto correlations. The total number of the latter fluctuations is almost a factor of ten greater than the total of the Hi-Mag fluctuations. Thus the number of active contributor fluctuations creating the correlations in the modified data sets after the replacements are made is almost a factor of ten smaller and these fluctuations represent a much smaller subset of the original data set. Naively one expects that the much smaller subset of originally active contributors to the correlations will either exhibit the original data set's correlations in a degraded form or show no correlations. Instead the Hi-Mag sentiment and price-change results both exhibit highly persistent auto correlations with the dominant ones in the Hi-Mag regimes being reversed from those in the original All-Mag sentiment and price-change data.

The reversal of the dominant correlations the Hi-Mag cases results in the probability of occurrence of a positive price-change to be smaller than that of a negative price-change producing a negative daily price-mobility skew excess. Similarly the probability of occurrence of a buyer trade is smaller than that of a seller one resulting in a negative daily order-flow skew excess. These negative price-mobility and order-flow intraday skews thus create the negative daily scale price-change and sentiment distribution skews at high magnitudes that are observed in the daily empirical distributions.

Section 3 first considered on the daily scale the positive and negative branches of the cumulative distribution of price-change and sentiment along with the behavior of the skew excesses with increasing fluctuation magnitudes. It found that the daily price-change and sentiment cumulative distributions for the two were nearly identical. The daily behaviors with increasing threshold magnitudes of the price-change and sentiment skew excesses were also very similar. The correlation between the latter two was 88%. This section then considered the same quantities on the seven-day scale. Here the price-change positive and negative cumulative distribution branches noticeably widened their separations both at low and at high magnitudes with the widening being larger at high magnitudes. This indicates appreciable overall skew amplification over their daily values. The sentiment's positive and negative branches of the cumulative distribution, on the other hand, were essentially the same as the ones on the daily scale. This section then compared the sentiment behaviors with increasing threshold magnitudes on the daily and seven-day scales. The comparison showed that the seven-day scale's negative sentiment skews systematically trended above the ones of the daily scale, indicating most of sentiment skew' magnitudes decrease below their daily values. For positive skews the seven-day skews showed various amplification/diminishment behaviors. The erratic amplifications and diminishments from daily values of the sentiment skews were due to the classification errors in the method used to identify buyer and seller trades.

The earlier study showed that intraday sign-correlations of price-mobility create the positive skew in the daily price change distribution. It argued that the positive daily skew then gives rise to the growth of skews with increasing scales in the following way. To form the two-day scale variable one adds two sequential daily variables. Since each of the individual days added together, on average contain the daily bias for positive skew, the process of addition amplifies at the two-day scale the initial positive daily bias, giving a stronger bias to positive skew at the two-day scale. Such amplification continues to act at each transition from scale s to scale $s + 1$. Hence the initial positive bias in the daily price-changes along with its amplification assures that the all-magnitude skews increase at successively higher scales.

The current study shows that this amplification process also operates at high magnitudes for price-change where the initial daily bias is negative. The additions of daily variables to form higher scale ones also amplifies the magnitudes of skews in the Hi-Mag regions of the distribution as shown by the large amplification factors over their daily values in Figure 3.5. In the absence of the error introduced by the buyer and seller classification method, one expects the above considerations would also apply to order-flow correlations' creation of positive All-Mag and negative Hi-Mag biases in the daily sentiment with the subsequent amplification of these biases at higher than daily scales.

Note that the present work and the earlier study do not address how the amplifications of daily skews with increasing scales arise from the successive addition of daily scale variables to form the progressively higher scale variables. These two studies simply report the empirical results showing such amplifications at the higher scales as obtained from the IVE data set. Such daily skew excess amplifications at higher scales are also found for price-change distributions in historical data sets of closing daily prices for the S&P 500, Dow Jones, and NASDAQ Indexes.

Closing this discussion, let's conjecture on the sources behind the negative skew excess at high magnitudes. The negative skew excesses at high magnitudes are suggestively in line with prospect theory of behavioral economics (cf. Kahneman 2011). Prospect theory holds that human behavior magnifies the fear of potential losses to a much greater degree than it does the attraction to potential gains. Thus it implies that agents initiating sell orders should act more quickly than those initiating buy ones, and there should be more active sellers than buyers when large magnitude price swings ensue. The quicker selling by a larger pool of sellers should give greater daily trade totals for sellers than those of buyers at high magnitudes. The fact that the stakes are bigger for high magnitude price-changes should enhance the trader's natural human fear of losses and cause them to sell more quickly in larger order lots that have more price impact. Thus prospect theory helps to explain the negative empirical skews of sentiment at high magnitudes. Additionally, regulations and business practice tend to encourage more rapid selling for large and sustained price moves since fund managers must issue quarterly reports on holdings. As noted by Keim and Madhavan (1995), once the decision to sell is made, the institutional trader is penalized more by selling a sinking stock too slowly than by buying a surging one too slowly. This is because the slow selling incurs a measurable accounting loss while the slow buying produces an unobservable opportunity loss. The above reasoning thus implies that the skews should be negative at the high magnitudes for both price-change and bull-bear sentiment.

The above paragraph's arguments give an intuitive understanding of why Hi-Mag skews should be negative. Corresponding intuitive arguments for positive All-Mag skews are elusive. Thus their positive empirical values present an interesting puzzle.

Appendix: Sign Auto Correlations of Skewed Independent Distributions

This appendix derives explicit individual positive and negative auto correlations as well as the average sign correlations at all lags for the case of independent fluctuations where the probabilities of positive and negative fluctuation are determined by the skew excess.

For independent fluctuations the probabilities of positive and negative fluctuations are given by $P_+ = \frac{1+\varepsilon}{2}$ and $P_+ = \frac{1-\varepsilon}{2}$, respectively, where ε denotes the skew excess[16] of the fluctuations at either the Hi-Mag or All-Mag regime. Using the sign function $sgn[\cdot]$ one writes the sign auto correlations as

$$C(sgn[z], sgn[z]), \tau) = \frac{1}{N_{tot}} \sum_{i=\tau}^{N_{tot}} sgn[z_i] sgn[z_{i-\tau}] \qquad \text{(A.1)}$$

Each fluctuation z_i is independent of all others and the probability of a matching positive (negative) fluctuation to follow an initial positive (negative) one is equal to $P_+P_+/\Sigma\ \ (P_-P_-/\Sigma)$, while the probability of a mismatched negative (positive) or positive (negative) fluctuation to follow an initial positive (negative) one or an initial negative (positive) one, respectively, is $2P_+P_-/\Sigma$ where the normalizing factor is $\Sigma = P_+P_+ + P_-P_- + 2P_+P_-$. The fluctuations are independent of each other, so one can express the sign auto correlations as

[16] This appendix presents the price-change skew excess and its auto correlations. The results obtained here also apply to skew excesses and correlations for skewed independent fluctuations of sentiment, price-mobility, and order-flow.

$$C(sgn[z], sgn[z], \tau) = \frac{[N_{tot}-\tau]\{P_+P_+ + P_-P_- - 2P_+P_-\}}{N_{tot}\Sigma} \cong \frac{\{(1+\varepsilon)^2+(1-\varepsilon)^2-2(1+\varepsilon)(1-\varepsilon)\}}{\{(1+\varepsilon)^2+(1-\varepsilon)^2+2(1+\varepsilon)(1-\varepsilon)\}} = \varepsilon^2 \qquad \text{(A.2)}$$

This equation assumes that $\tau/N_{tot} \ll 1$ so that $[N_{tot} - \tau]/N_{tot} \cong 1$.

Now consider the individual correlations $C_{++}(\tau)$ and $C_{--}(\tau)$. In terms of these correlations the probabilities of a sign matching fluctuation to follow the initial one are given by $\frac{1+C_{++}(\tau)}{4}$ and $\frac{1+C_{--}(\tau)}{4}$ for positive and negative fluctuations, respectively. Equating each of the latter respective probabilities to the corresponding ones in equation (A1.2) above and noting that the denominator of the fraction following the approximate equality in this equation sums to four, one writes the following relations for the individual auto correlations

$$\frac{1+C_{++}(\tau)}{4} = \frac{(1+\varepsilon)^2}{4} \to C_{++}(\tau) = 2\varepsilon + \varepsilon^2 \text{ and } \frac{1+C_{--}(\tau)}{4} = \frac{(1-\varepsilon)^2}{4} \to C_{--}(\tau) = -2\varepsilon + \varepsilon^2. \qquad \text{(A.3)}$$

Hence one obtains $\bar{C}(\tau) = \frac{1}{2}[C_{++}(\tau) + C_{--}(\tau)] = \varepsilon^2 = C(sgn[z], sgn[z], \tau)$. (A.4)

REFERENCES

Bouchaud, J-P., Farmer, J. D., Lillo, F., 2009, "How Markets Slowly Digest Changes in Supply and Demand," Handbook of Financial Markets: Dynamics and Evolution, p. 57

Bouchaud, J-P., Gefen, Y., Potters, M., and Wyart, M., 2004, "Fluctuations and response in financial markets: The subtle nature of random price changes", Quantitative Finance, Vol. 4(2), p. 176

Kahneman, D., 2011, "Thinking, Fast and Slow" (Farrar, Straus and Giroux, New York 2011)

Keim, D. B., Madhavan, A, 1995, "Anatomy of the Trading Process: Empirical Evidence on the Behavior of Institutional Traders," Journal of Financial Economics, Vol.37, p. 371

Lee, C., Ready, M., 1991, “Inferring trade direction from intraday data,” Journal of Finance 46, p. 733

Lillo, F., and J. D. Farmer, 2004, "The long memory of the efficient market", Studies in Nonlinear Dynamics & Econometrics Vol. 8(3), p. 1

Mazurek, T, J., 2017,"Persistent Intraday Correlations Create Skews in Daily-Scale Distributions," The link to the paper is here.