Over- and Under-Reaction in Liquid Markets

Alexei Chekhlov
Systematic Alpha Management, LLC; Partner & Portfolio Manager
Columbia University; Visiting Assistant Professor

September 29, 2009

(submitted to The Hedge Fund Journal, England)

It is currently believed that the so-called "Behavioral Finance" (BF) plays a significant role in explaining the existing irrational behavior of investors, including that of professional investors. Although there exist some notably good examples that seem to illustrate this point of view (such as the recent book by Dan Ariely[1], and more classical collections by Richard H. Thaler[2], to name a few), except for a few toy-models, they seem to provide little, if any, practical help in looking for traces of irrational behavior in actual financial price series. Here we would like to present some direct quantitative statistical testing over raw financial data which gives BF some practical justification. Let us assume that if the financial world has consisted of only the rational agents, then the series of price changes of any financial assets would follow a Random Walk (RW). Then, the question of looking for inefficiencies is equivalent to studying the differences in statistics of a given financial asset price changes (or returns) price series to that of a RW. Various statistical measurement techniques may allow one to register and quantify such deviations from RW as existence of: serial (negative or positive) autocorrelation in price changes, long memory effects, "fat" tails of probability density functions, and multi-scaling or intermittency effects. Here we will be concerned with the purely phenomenological question of existence of inefficiency, rather than a more complex question of whether it "survives" some transaction costs for a particular trading strategy.

First we would like to look at the price changes of the "best known" and most liquid index in the world, S&P 500, traded through E-Mini futures. Due to its ample liquidity this instrument is widely perceived as very efficient. To analyze the agreement with the RW model, we have considered two simple experiments for which we have used 1-minute back-adjusted futures data since 9/10/1997 (inception) until 8/31/09.

For a range of time separations τ used for price change calculation, we would like to compare $N_+(\tau)$, the number of two consecutive price movements in the same direction (sequence), with $N_-(\tau)$, the number of two consecutive price movements in the opposite direction (reversals). According to RW, the ratio $N_-(\tau)/N_+(\tau)$ should be independent of τ and equal to 1 within some noise. The results of actual measurements are presented in Fig. 1. We see that for τ>380 minutes, the agreement with the RW relationship N_-

[1] Dan Ariely, "Predictably Irrational", Revised and Expanded Edition: The Hidden Forces That Shape Our Decisions, Harper, 2009.
[2] Richard H. Thaler, "The Winner's Curse", Paradoxes and Anomalies of Economic Life, Princeton University Press, 1994; "Advances in Behavioral Finance", Vol. I and II, Princeton University Press, 2005.

$(\tau)/N_+(\tau)=1$ is quite good indeed. For $1 \le \tau < 114$ minutes however, we observe a statistically strong (about 16% at the peak for $\tau=4$) excess of reversals over the sequences. This is a clear deviation from RW in the direction of *short-term over-reaction* of high-frequency price changes in this market. The agents over-react to the news, later correcting their own actions. For $115 \le \tau < 380$ minutes approximately, there is evidence for a weaker (around 5% at the peak for $\tau=114$) excess of sequences over the reversals. The total sample size in this and next experiment was over 1.2 million.

In the second experiment, for every length of one-directional excursion of price we will measure the frequencies of such excursions and we will compare them to the similar frequencies obtained in a RW model. We will do so by replacing every 1-minute price change $\Delta p=p(t+1)-p(t)>0$ with +1, with -1 if $\Delta p<0$, and for every $\Delta p=0$ the sign of the previous step will be extended. This way we will break the overall time series of price changes into chains of consecutive +1's and -1's, and we will be concerned with the number of same sign chains $N_+(k)$ or $N_-(k)$ for every particular length k. It is easy to see that for a RW of length M with the probability of upward 1-minute move $p=N_+(1)/M$, these two numbers should be equal to: $N_+(k)=(M-k+1)\cdot p^k$, and $N_-(k)=(M-k+1)\cdot(1-p)^k$. The results in Fig. 2 are shown for $1 \le k \le 16$ in the form of relative excess number to RW values: $(N_{Actual}-N_{RW})/ N_{RW}$, for all cases with sample size larger than 100 (shown on the secondary y-axis). It is obvious from Fig. 2 that even though for k=1 the agreement with the RW is perfect. For k>1 two things happen: first, the excess increases from k=2 symmetrically up to k=12, then the positive chains start to outnumber the negative chains. The excess above RW count is very statistically significantly >0, exceeding 600% for k=16. This is a clear deviation from RW in the direction of *under-reaction* in price changes for large enough time-separations and, possibly, some weak asymmetry between positive and negative price changes. Here, new information coming into the market, instead of being absorbed into the price instantly in accordance with the RW, is absorbed over finite time during which the trading agents are watching each other's actions, waiting for additional self-reassurance in seeing others do the same.

Now, let us look at the full auto-correlations for the high-frequency price changes in such liquid instruments, as equity indices and currency futures. We will do so using the "variance ratio"[3] test of discrete price changes on equidistant time (again, 1-minute):

$$VR(q) = 1 + 2 \cdot \sum_{k=1}^{q-1} \left(1 - \frac{k}{q}\right) \cdot \rho_k,$$

where ρ_k is autocorrelation coefficient of two price changes separated by k time intervals. This formula for any time separation q explicitly measures the ratio of the actual variance to the RW variance. In Fig. 3 and 4 we have measured VR(q) for returns of day-session price data in a set of liquid futures markets belonging to two classes: equity indices (S&P 500, Dow Jones, NASDAQ-100, FTSE-100, CAC-40, DAX, DJ-Euro Stoxx) and currencies (Euro, British Pound, Swiss Franc, Japanese Yen, Canadian Dollar, Australian Dollar) since inception of each of these markets until 08/31/2009. With the exception of two markets of Euro and British Pound, for all other instruments, VR(q) is reasonably

---

[3] John Y. Campbell, Andrew W. Lo, A. Craig MacKinlay, "The Econometrics of Financial Markets", Princeton University Press, 1997.

convex in the high-frequency area, experiencing most of its curvature for 1[minute]$<\tau\leq$2[trading days], which corresponds to strong deviations from the RW. For $\tau>$2[trading days] the VR(q) is quite flat and the agreement with the RW is reasonably good. For longer than a day time separations, Euro looks to be in good agreement with the RW, whereas British Pound looks like having some under-reaction or trend-following properties. In reality, the above trend-following properties of British Pound are heavily dominated by the early history period of 1977-1990, which is simply absent from the Euro history. For short time separations all currencies seem to exhibit over-reaction. If VR(q) was re-measured during 1990-present, nearly perfect agreement with the RW will be established. In difference to the indices, here more liquidity seems to indeed be related to more efficiency, which could be related to the fact that there is much less retail human participation in the currency markets. Therefore, for both of these most liquid futures cases, there is evidence for quite universal (across different countries) statistically strong short-term (within 2 trading days) *over-reaction* or mean-reversion. A more advanced reader may note that some deviations from VR(q)=1 may be due to the so-called "fat tails" of the probability density function of price differences. Indeed, this could be a valid note; however, the above over-reaction survives even after adjusting for this phenomenon.

As the above examples involve some of the most liquid futures instruments in the world, this, hopefully, illustrates that *high instrument liquidity is not necessarily synonymous with high market efficiency.*

The last example we would like to give here is from the area of stock prices reacting to the positive and negative Earnings-Per-Share (EPS) announcements. In this example the frequency of the price data is daily, the accounting variables – quarterly. The database consists of full survivorship-bias-free price and fundamental information, adjusted for corporate events for all North American public equity issuers from 1/1993 to 8/2009. Leaving out numerous details[4], we have mechanically defined "positive" and "negative" EPS surprise filters in order to see how the average price changes for up to 50 trading days after an EPS surprise. "Positive surprise" stocks will be bought, and "negative surprise" stocks will be sold short at the next open of the following the announcement day (time t). Statistically, for these two separate groups of cases and across all such surprise events we would calculate the average price change conditional on either "positive" or "negative" surprise: $R_+(\tau) = \langle p(t+\tau)-p(t)|\text{"positive surprise"}\rangle$ and $R_-(\tau) = \langle p(t)-p(t+\tau)|\text{"negative surprise"}\rangle$. These functions are depicted in Fig. 5 and 6 in their absolute compounded, and market-neutral (with the S&P 500 return subtracted) compounded versions. For the case of the "positive surprise" filter (Fig. 5) we observe a statistically significant *under-reaction* to positive news that decays in its market-neutral version after about 15 trading days. This is evidence of a delayed reaction among the agents that transact in such stocks when they are afraid to instantly price-in the news and are looking at each other in search of confirmation for their actions, increasing their

---

[4] "Positive surprise" we call a stock that after announcing its accounting data satisfies: EPS>0, positive trail of earnings: EPS(-1)>0, EPS(-2)>0, EPS(-3)>0, EPS(-4)>0 – for earnings per share 1, 2, 3, and 4 quarters ago, positive surprise definition: if Surprise=EPS-EPS(-1)/Abs[EPS(-1)], then Surprise>10%. "Negative surprise" we call a stock that has: EPS(-1)<-0.2, Surprise<25%.

position as they find the confirmation and thus, moving the price further up later. For the case of the "negative surprise" filter (Fig. 6) we observe a case of short-term over-reaction that decays to zero in its market-neutral version after 5 trading days or so. Here, in emotional reaction to the news the agents oversell the stock, after 5 days somewhat correcting their opinion (what is sometimes called "dead cat bounce" effect). Although the statistical accuracy of both these results is noticeably smaller than for the above high-frequency cases, we do not find any bias of these results to a particular bull or bear markets or other sub-sampling.

To summarize, properly constructed statistical tests of the financial data even for the most liquid financial instruments can reveal sufficient evidence for both over- and under-reaction, over-reaction concentrated to shorter time scales. Such and similar tests bridge the gap between the rational scientific view and BF view on quantitative finance. If, according to BF, the source of over- and under-reaction is hidden in human-like properties of the trading agents, then the best way to exploit these inefficiencies is to trade via automated "robots" or algorithms specifically designed to trade against such human-inflicted market inefficiencies as long as they exist. "Robots" do not need sleep, they are emotionally agnostic, they are free from "anchoring", and they will relentlessly pursue the objectives built into them. Ironically, their Achilles' heel is as the perfectly rational "robots" will dominate over the human-like traders, the above inefficiencies will cease to exist.

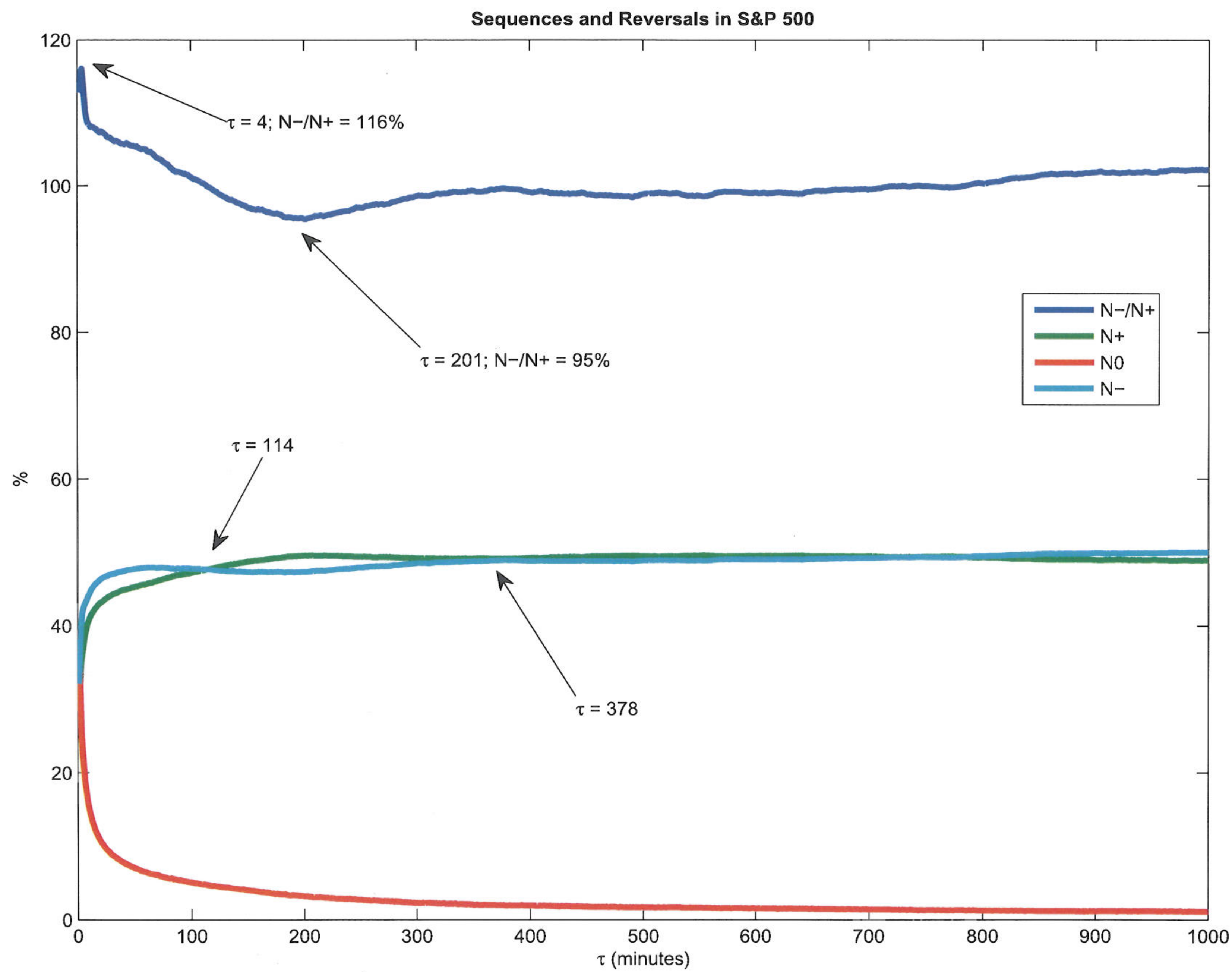
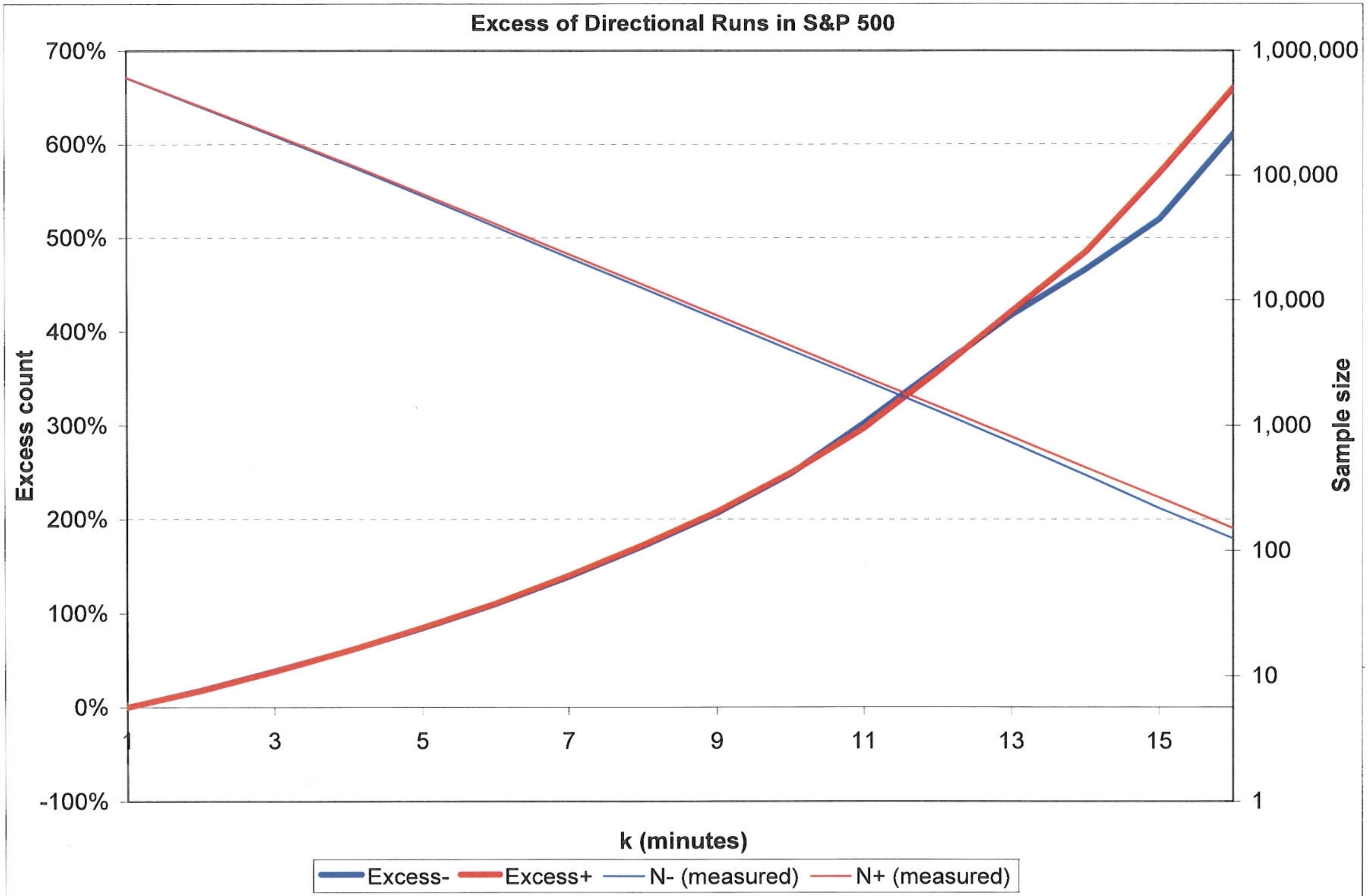
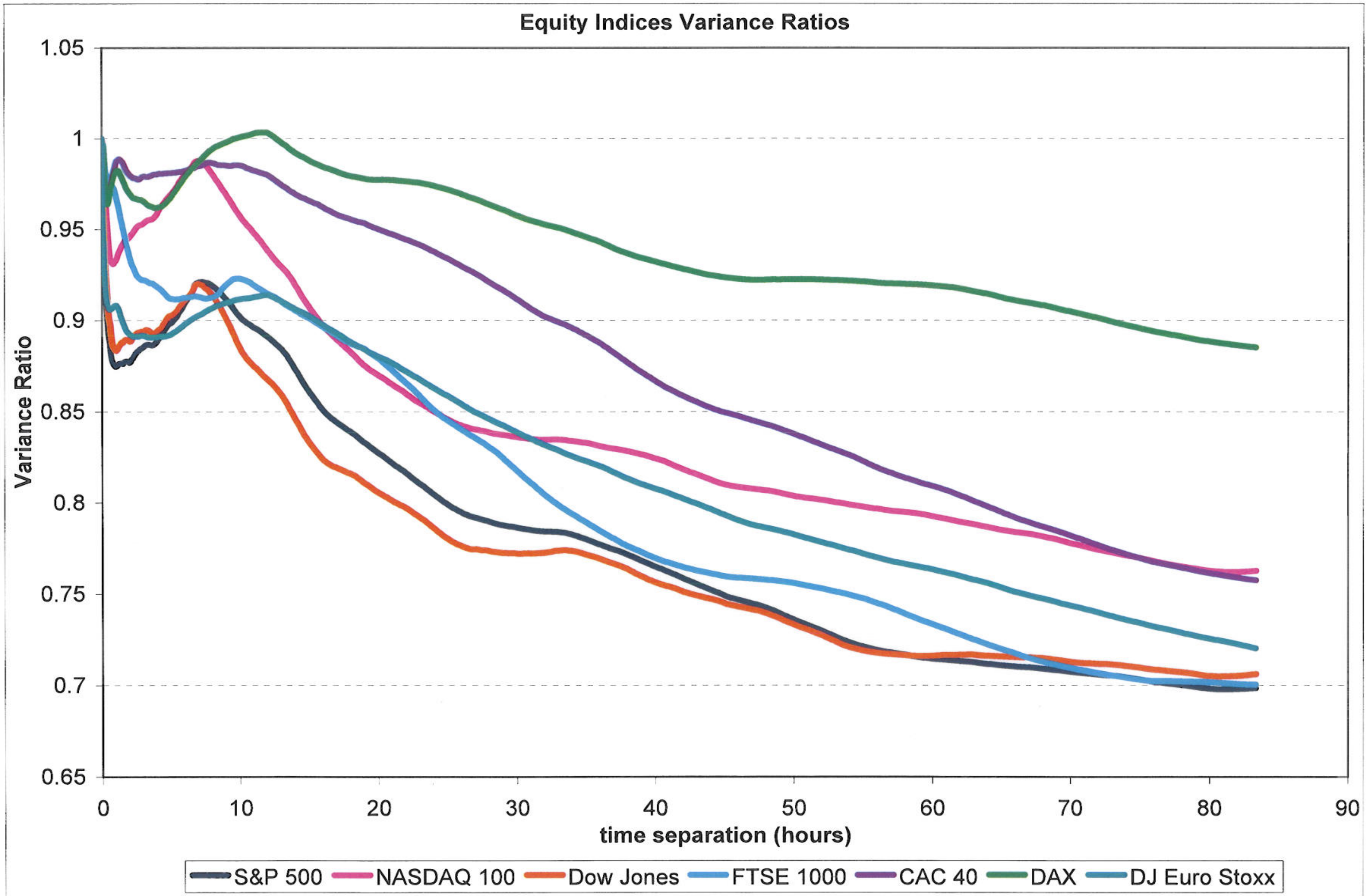

Sequences and Reversals in S&P 500
120
100
80
60
40
20
0
%
0
100
200
300
400
500
600
700
800
900
1000
τ (minutes)
τ = 4; N−/N+ = 116%
τ = 201; N−/N+ = 95%
τ = 114
τ = 378
N−/N+
N+
N0
N−

Fig. 1



Excess of Directional Runs in S&P 500
700%
600%
500%
400%
300%
200%
100%
0%
-100%
Excess count
1,000,000
100,000
10,000
1,000
100
10
1
Sample size
1
3
5
7
9
11
13
15
k (minutes)
Excess-
Excess+
N- (measured)
N+ (measured)

Fig. 2.



Equity Indices Variance Ratios
Variance Ratio
1.05
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0
10
20
30
40
50
60
70
80
90
time separation (hours)
S&P 500
NASDAQ 100
Dow Jones
FTSE 1000
CAC 40
DAX
DJ Euro Stoxx

Fig. 3

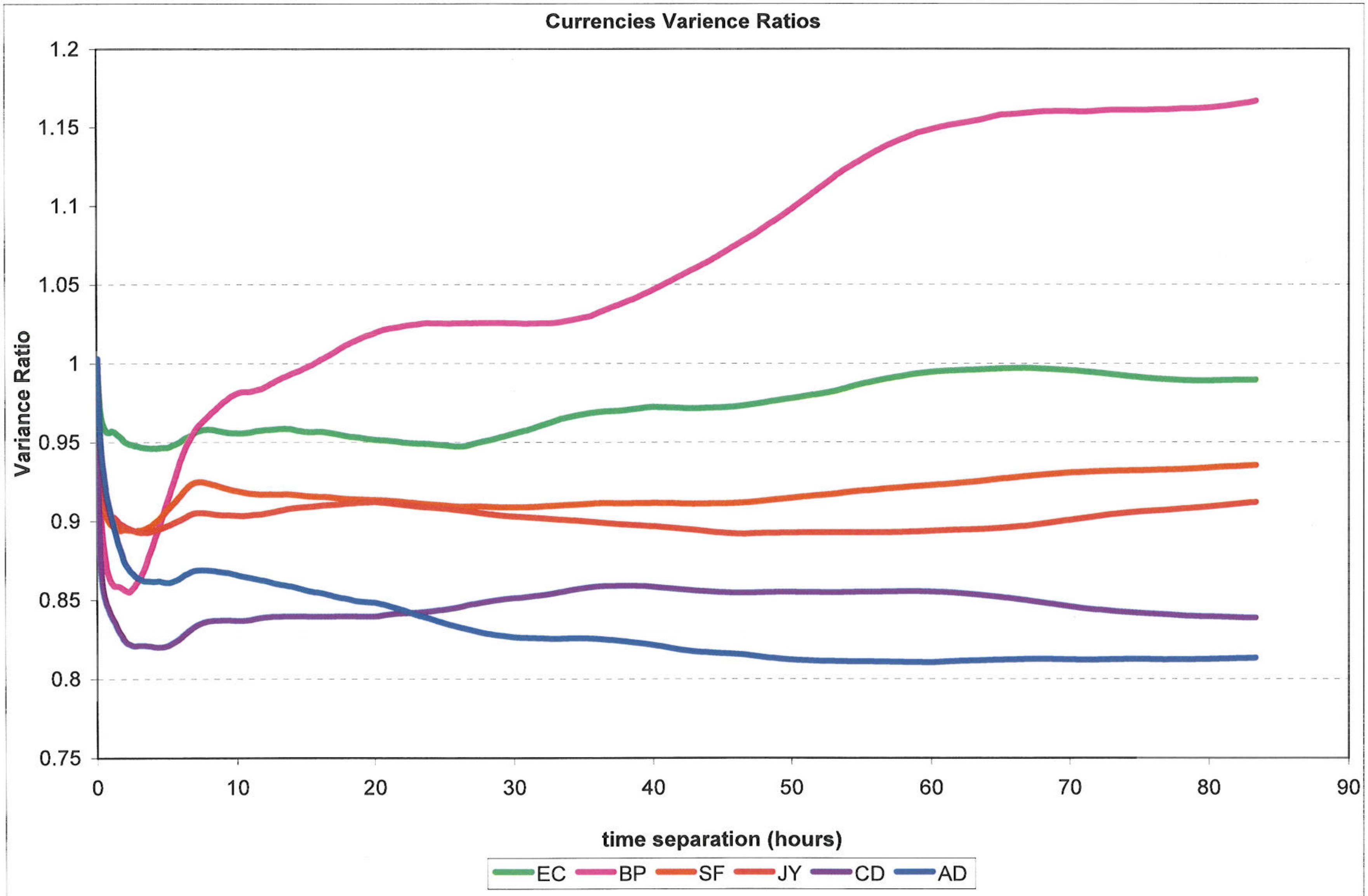
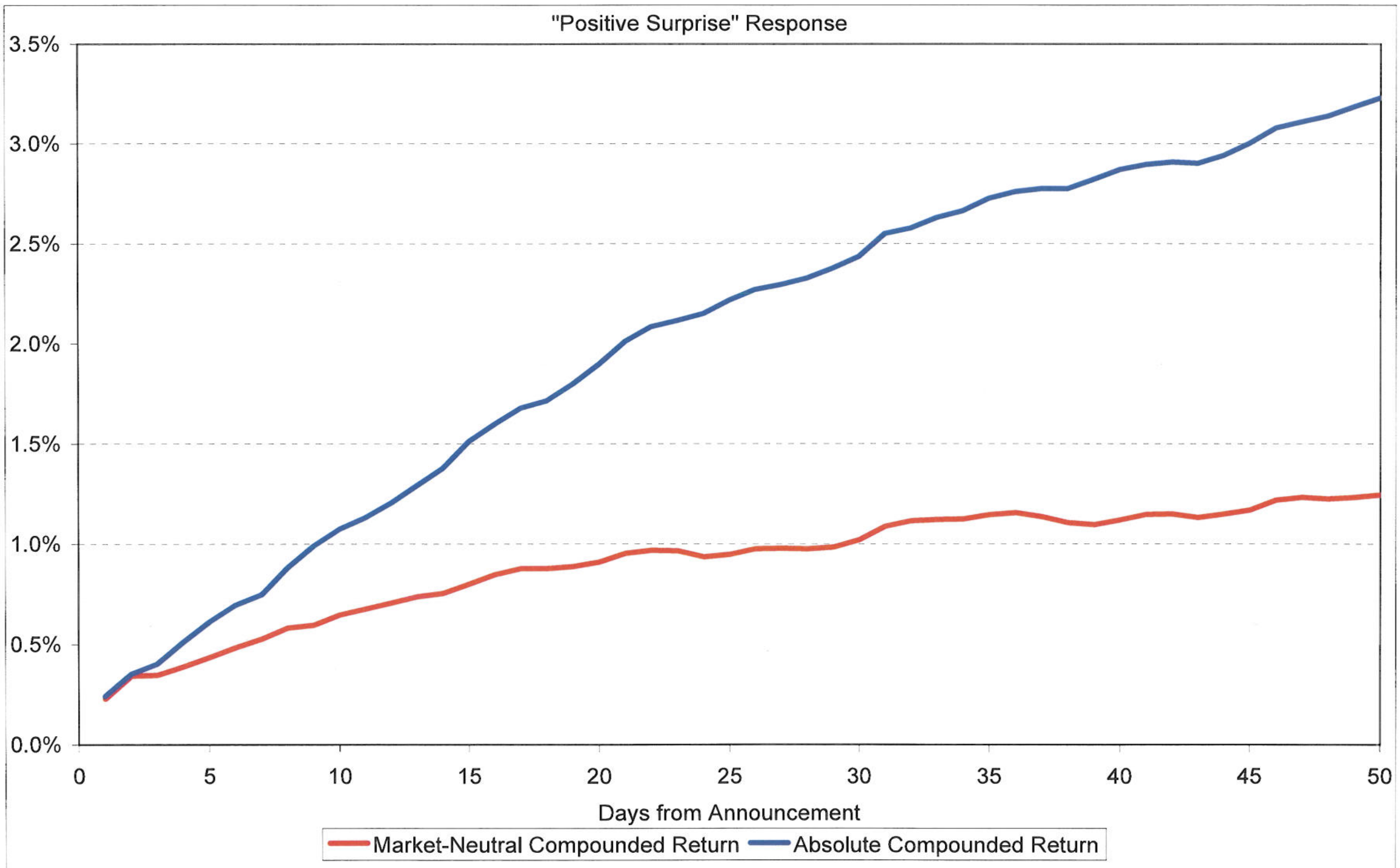

Currencies Varience Ratios
Variance Ratio
1.2
1.15
1.1
1.05
1
0.95
0.9
0.85
0.8
0.75
0
10
20
30
40
50
60
70
80
90
time separation (hours)
EC
BP
SF
JY
CD
AD

Fig. 4.



"Positive Surprise" Response
3.5%
3.0%
2.5%
2.0%
1.5%
1.0%
0.5%
0.0%
0
5
10
15
20
25
30
35
40
45
50
Days from Announcement
Market-Neutral Compounded Return
Absolute Compounded Return

Fig. 5.

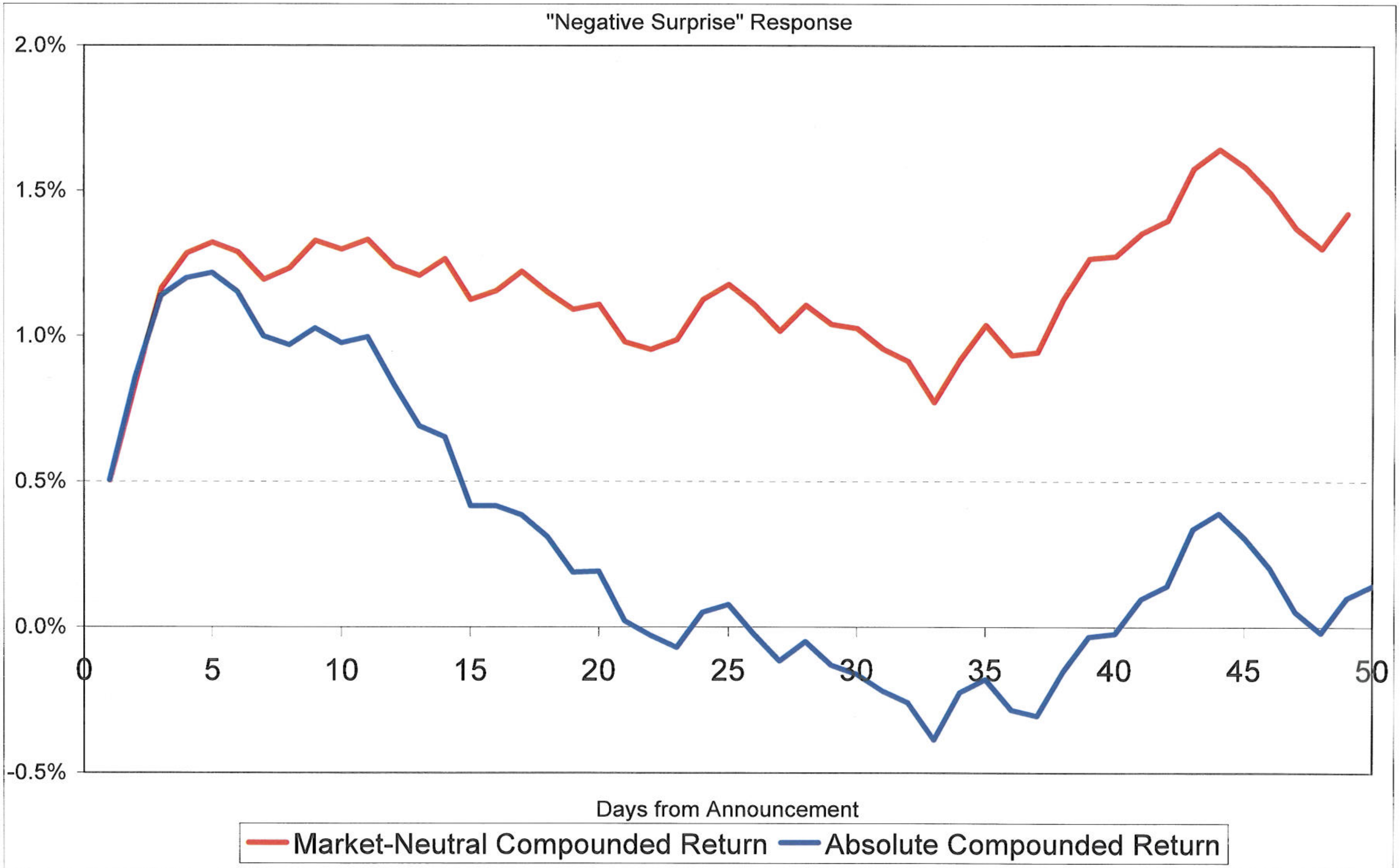

"Negative Surprise" Response
2.0%
1.5%
1.0%
0.5%
0.0%
-0.5%
0
5
10
15
20
25
30
35
40
45
50
Days from Announcement
Market-Neutral Compounded Return
Absolute Compounded Return

Fig. 6