Math Methods – Financial Price Analysis

Mathematics GR5360

Instructor: Alexei Chekhlov

Stationarity, Correlation and Memory*

Let us discuss the property of stationarity of price changes, first. When the stochastic variables are independent, stationarity implies that the stochastic process $x(t)$ is independent identically distributed (iid). The statistical observables (measurables) characterizing a stochastic process can be written in terms of n-th order statistical properties. For example, in the 1-st order case it is sufficient to define the mean:

$$\langle x(t) \rangle = \int_{-\infty}^{\infty} x \cdot P(x,t) \cdot dx,$$

where $P(x,t)$ is the probability density function of observing the random variable x at time t. In the 2-nd order case, it is sufficient to define the auto-correlation function:

$$\langle x(t_1) \cdot x(t_2) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 \cdot x_2 \cdot P(x_1, x_2; t_1, t_2) \cdot dx_1 \cdot dx_2,$$

where $P(x_1, x_2; t_1, t_2)$ is the joint probability density function of observing the random variable x_1 at time t_1 and the random variable x_2 at time t_2.

Although to fully describe the statistical properties of a stochastic process, we need to know the joint probability density function $P(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n)$ for every x_i, t_i and n, but in most studies it is sufficient to know only the two-point statistics.

* - with some changes from "An Introduction to Econophysics" by Mantegna and Stanley, ref. B1.

Stationarity, Correlation and Memory*

In a strict sense, a stochastic process $x(t)$ is called stationary if its probability density function $P(x,t)$ is invariant under a time shift: $t \to t + \Delta t$. Some less-restrictive definitions of stationarity also exist. For example, a wide-sense stationary process is defined if the following three conditions are met:

$$\langle x(t) \rangle = \mu,$$

$$\langle x(t_1) \cdot x(t_2) \rangle = R(t_1, t_2),$$

where $R(t_1, t_2) = R(\tau)$ is only a function of $\tau = t_2 - t_1$, and

$$\langle x^2(t) \rangle = R(0).$$

Therefore, the variance of such process, $R(0) - \mu^2$, is time-independent.

* - with some changes from "An Introduction to Econophysics" by Mantegna and Stanley, ref. B1.

Stationarity, Correlation and Memory*

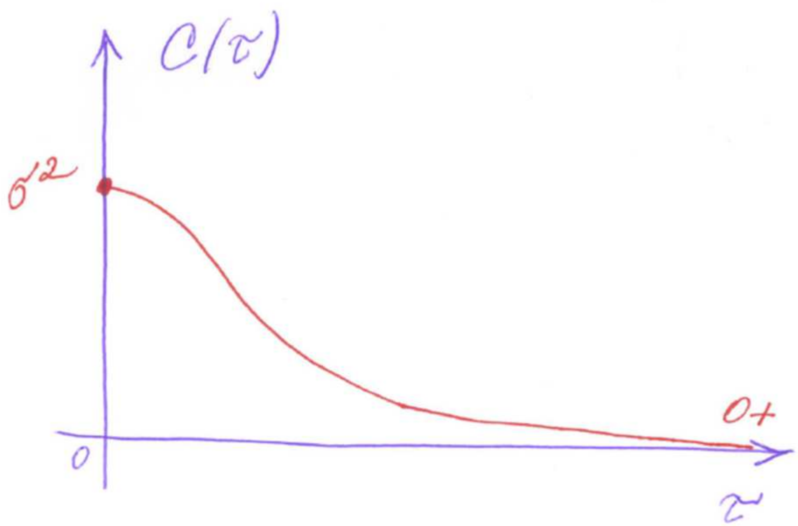
The auto-correlation function introduced above is dependent on the average value of the stochastic process. If the average value is different from zero, one needs to accomodate this by considering auto-covariance:

$C(t_1, t_2) \equiv R(t_1, t_2) - \mu(t_1) \cdot \mu(t_2).$

For stationary processes the auto-covariance is equal to:

$C(\tau) = R(\tau) - \mu^2.$

For a positively correlated stochastic process the typical shape of the auto-correlation function will be: it starts at $C(0) = \sigma^2$, and continuously falls off to $C(\tau) \to 0+$ for $\tau \to +\infty$. For simplicity we can consider stochastic processes with zero mean $\mu = 0$, and unit variance: $R(0) = \sigma^2 = 1$.



* - with some changes from “An Introduction to Econophysics” by Mantegna and Stanley, ref. B1.

Stationarity, Correlation and Memory*

The physical meaning of the auto-correlation has to do with the so-called (time) "memory" of stochastic process. If we consider the integral of the dimensionless auto-correlation function $R(\tau)$ for a stationary process, then the area below $R(\tau)$ has the dimension of time (a typical time-scale or time "memory"):

$$\tau_c = \int_0^{+\infty} R(\tau)d\tau = \begin{cases} \text{finite - "short" memory;} \\ \text{infinite - "long" memory.} \end{cases}$$

This integral has the physical dimension of time. When this integral is finite, there exists a typical time memory τ_c called the correlation time of the process.

Some notable examples:

1) $R(\tau) = e^{-\tau/\tau_c}$: $\int_0^{\infty} \mathrm{e}^{-\tau/\tau_c} d\tau = \tau_c \cdot \int_0^{\infty} e^{-y} dy = \tau_c \cdot \left(e^0 - e^{-\infty}\right) = \tau_c.$

2) $R(\tau) = e^{-\tau^{\nu}/\tau_0}$: $\int_0^{\infty} e^{-\tau^{\nu}/\tau_0} d\tau = \frac{\tau_0^{1/\nu}}{\nu} \cdot \Gamma\left(\frac{1}{\nu}\right).$

3) $\mathrm{R}(\tau) \propto \tau^{\eta-1}$: for $0 < \eta \le 1$, $\int_{t_1}^{\infty} \tau^{\eta-1} d\tau = \infty.$

* - with some changes from “An Introduction to Econophysics” by Mantegna and Stanley, ref. B1.

Stationarity, Correlation and Memory*

The finiteness of the area under the auto-correlation function carries information about the typical time scale of the memory of the process. In fact, as a zero-order approximation, it is possible to model the system by assuming that the full (= +1) correlation is present from the smallest up to scales τ^*, and no correlation (= 0) is present for $\tau > \tau^*$, where τ^* is the area under the auto-correlation function. Obviously, not all integrals of the monotonically decreasing functions are finite!

For example, in case (3) above, it is impossible to select a time scale that can separate a regime of temporal correlations from a regime of pairwise independence. Random variables characterized by an auto-correlation function such as in case (3) are called to be long-range correlated.

* - with some changes from "An Introduction to Econophysics" by Mantegna and Stanley, ref. B1.

Stationarity, Correlation and Memory*

Let $S_n = \sum_{i=1}^{n} x_n$.

Then, grouping main diagonal and off-diagonal elements with the use of symmetry,

$$S_n^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j = \sum_{i=1}^{n} x_i^2 + 2\sum_{i=1}^{n}\sum_{k=1}^{n-i} x_i x_{i+k} .$$

With averaging, and using the notations $\sigma^2 = \langle x_i^2 \rangle$, and $c_k = \langle x_i x_{i+k} \rangle$, we get:

$$\langle S_n^2 \rangle = n\sigma^2 + \sum_{k=1}^{n-1} (n-k) c_k .$$

For $n \to +\infty$, we get:

$$\langle S_n^2 \rangle \to n\left(\sigma^2 + \sum_{k=1}^{n-1} c_k \right).$$

Depending on the behavior of the second term in this expression, we can distinguish two types of behavior:

I. The sum of correlation terms is finite for large n : $\lim_{n\to+\infty} \sum_{k=1}^{n-1} c_k < \infty$. In this case the random variables are weakly-dependent, or short-range-correlated, and then for large values on n we have: $\langle S_n^2 \rangle \propto n$.

* - with some changes from “An Introduction to Econophysics” by Mantegna and Stanley, ref. B1.

Stationarity, Correlation and Memory*

2. The sum of correlation terms diverges, or :

$$\lim_{n\to+\infty}\sum_{i=1}^{n}\sum_{k=1}^{n}\langle x_i x_{i+k}\rangle=\infty.$$

In this case the random variables are said to be strongly dependent or long - range correlated. Long - range correlated random variables are characterized by the lack of a typical temporal scale. This behavior is observed in stochastic processes characterized by a power - law auto - correlation function, as in case (3) above.

Let us recall a very simple identity :

$$P(t_n)=P(t_0)+\sum_{i=1}^{n}\Delta P_i=P(t_0)+\sum_{i=1}^{n}\left(P(t_{i+1})-P(t_i)\right),$$

where $t_i=i\cdot\Delta t$, therefore, for a financial price series the integral of the auto - correlation function can be used to distinguish between processes with short - range and long - range memory.

* - with some changes from "An Introduction to Econophysics" by Mantegna and Stanley, ref. B1.

Short-Range Memory Random Processes*

We have noted that short-range correlated random processes are characterized by a typical time scale or finite memory. One example was already considered by us, namely with exponentially decaying auto-correlation function:

$$R(\tau) = \sigma^2 \cdot e^{-|\tau|/\tau_c}.$$

This auto-correlation function, for example, describes the statistical memory of the velocity $v(t)$ of a Brownian particle, as the auto-correlation function of $v(t)$. Such processes are "easy", as one can assume $R(\tau) = 0$ for $\tau > \tau_c$.

In addition to the characterization of the two-point statistical properties in terms of auto-correlation function, we can now investigate the same statistical properties in the Fourier-frequency domain. Recalling the Khinchin Theorem, the energy spectrum of a wide-sense stationary random process is the Fourier transform of its auto-correlation function:

$$E(\omega) = \int_{\tau=-\infty}^{+\infty} R(\tau) \cdot e^{-2\pi i \omega \tau} \cdot d\tau.$$

For the auto-correlation function above this integral is doable in closed form:

$$E(\omega) = \frac{2\sigma^2 \tau_c}{1 + (2\pi\omega\tau_c)^2}.$$

This energy spectrum has two asymptotical behaviors:

$$E(\omega) \to \begin{cases} 2\sigma^2\tau_c, \text{ for } \omega << \dfrac{1}{(2\pi\tau_c)}, \text{ which is frequency-independent "white noise";} \\ \dfrac{\sigma^2}{2\pi^2\tau_c} \cdot \dfrac{1}{\omega^2}, \text{ for } \omega >> \dfrac{1}{(2\pi\tau_c)}, \text{ which is Random Walk.} \end{cases}$$

To summarize, short-range correlated stochastic process can be characterized with respect to their statistical properties by investigating either the auto-correlation function or the energy spectrum. The fast-decaying auto-correlation functions and energy spectra resembling $\dfrac{1}{\omega^2}$ are the fingerprints of short-range auto-correlated stochastic processes with memory.

* - with some changes from "An Introduction to Econophysics" by Mantegna and Stanley, ref. B1.

Long-Range Memory Random Processes*

We have already seen that stochastic processes characterized by a power - law auto - correlation function are long - range correlated. Power - law auto - correlation functions are observed in many systems - physical, biological, and economic.

Let us consider a stochastic process with a energy spectrum of the form :

$$E(\omega) = \frac{C}{|\omega|^{\eta}}, \text{ where } 0 < \eta < 2 \text{ and } C = const.$$

On the previous page we concluded that the case $\eta = 0$ corresponds to white noise and the case $\eta = 2$ corresponds to Random Walk.

All the in - between cases $0 < \eta < 2$ are very interesting and are often called " $1/\omega^{\eta}$ - noise".

Such processes were observed in wide variety of phenomena : current fluctuations in diodes and transistors; fluctuations of velocity in strong turbulence; fluctuations in traffic low on a highway, etc.

Such processes are non - stationary. It is diffucult to distinguish $1/\omega$ - noise which has no characteristic time scale from a process with many characteristic time scales. These are the most interesting processes which in the case of financial applications do allow for some forecasting.

* - with some changes from “An Introduction to Econophysics” by Mantegna and Stanley, ref. B1.

Influence of Mean-Reversion on Variance

Earlier we have already introduced and simulated the effect of mean-reversion as modeled by an Ornstein-Uhlenbeck process. In what follows we will now consider the details of such mean-reverting processes, by solving first a discrete and then continuous version of it. Let us recall that for a Random Walk we have:

$V(\tau) \equiv \overline{(\Delta p)^2} = \overline{\left(p(t+\tau) - p(t)\right)^2} \propto \tau$, or $S(\tau) = \sqrt{V(\tau)} \propto \sqrt{\tau}$.

Now, let us consider the following discrete mean-reverting process:

$$\begin{cases} x_{k+1} = \alpha \cdot x_k + \xi_k, \text{ for } 0 < \alpha \leq 1, \text{ and } \xi_k \text{ - i.i.d. random variables;} \\ x_0 = 0. \end{cases}$$

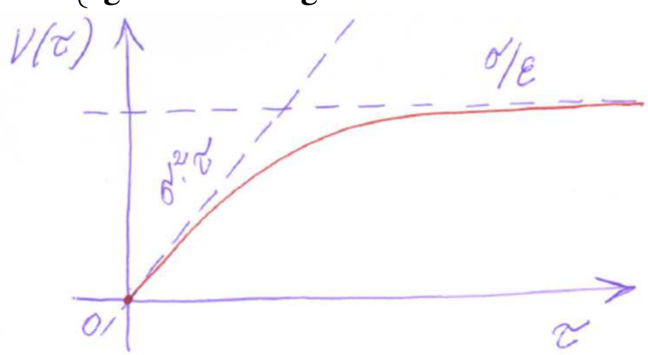
The solution (by recursive substitution) of it is: $x_k = \sum_{l=0}^{k-1} \alpha^{k-l-1} \cdot \xi_l$. Using it, one can show that for variance:

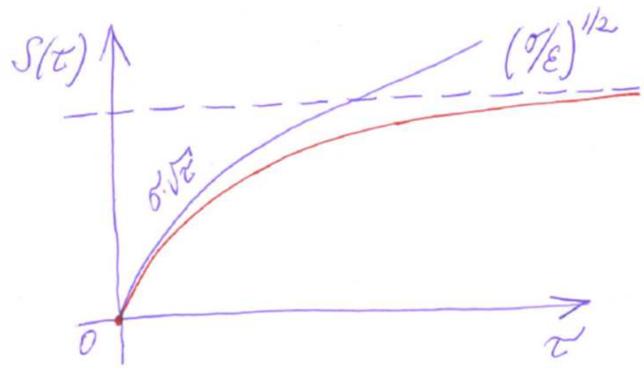
$$V(\tau) = 2\sigma^2 \cdot \frac{1-\alpha^\tau}{1-\alpha^2} \to \frac{\sigma^2}{\varepsilon} \cdot \left(1 - e^{-\varepsilon\tau}\right) \text{if } \alpha = 1 - \varepsilon \text{ for } 0 < \varepsilon << 1.$$

Therefore we can consider two limiting cases:

$$V(\tau) \to \begin{cases} \sigma^2\tau, \text{ for } \tau << \frac{1}{\varepsilon}, \text{ which corresponds to Random Walk dominating;} \\ \frac{\sigma^2}{\varepsilon}, \text{ for } \tau >> \frac{1}{\varepsilon}, \text{ which corresponds to mean-reversion dominating.} \end{cases}$$

In graphical form:





Influence of Mean-Reversion on Variance

Now let us proceed with the continuous linear Ornstein - Uhlenbeck model case:

$$dx = -\alpha \cdot (x - \bar{x}) \cdot dt + \sigma \cdot dW.$$

Let us seek the solution in the form: $x = y \cdot e^{-\alpha t}$. Then after substitution we get:

$dy \cdot e^{-\alpha t} = \alpha \cdot \bar{x} \cdot dt + \sigma \cdot dW$, which has the general solution:

$$y(t) = y(0) + \bar{x} \cdot \left(e^{\alpha t} - 1\right) + \sigma \cdot \int_{s=0}^{t} e^{\alpha s} \cdot dW(s).$$

After reverting to $x(t)$ we get:

$$x(t) = x(0) \cdot e^{-\alpha t} + \bar{x} \cdot \left(1 - e^{-\alpha t}\right) + \sigma \cdot \int_{s=0}^{t} e^{\alpha(s-t)} \cdot dW(s).$$

Using this general solution we can see that for the mean we have:

$$\langle x(t) \rangle \to \bar{x}, \text{ if } t \to +\infty.$$

For the covariance we have:

$$\left\langle \left(x(t) - \bar{x}\right)\left(x(s) - \bar{x}\right) \right\rangle = \frac{\sigma^2}{2\alpha} \cdot e^{-\alpha(s+t)} \cdot \left(e^{2\alpha \cdot \min(s,t)} - 1\right).$$

Lastly, for the variance we have:

$$\left\langle \left(x(t) - \bar{x}\right)^2 \right\rangle = \frac{\sigma^2}{2\alpha} \cdot e^{-2\alpha t} \cdot \left(e^{2\alpha t} - 1\right) \to \frac{\sigma^2}{2\alpha}, \text{ if } t \to +\infty.$$

Andrew Lo's Variance Ratio Test

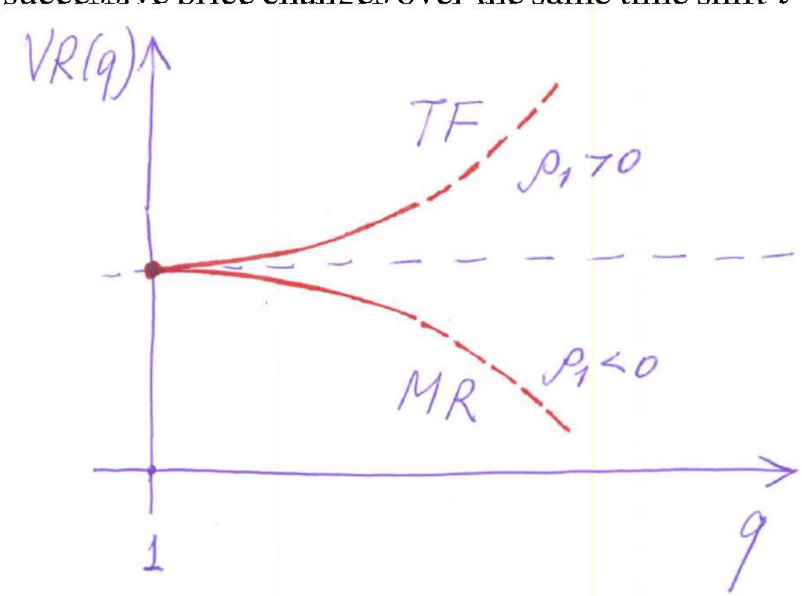
A very simple test was proposed by Andrew Lo et. al. around 1987 to compare statistical properties of a random time series to those of a Random Walk. Let us consider two successive price changes over the same time shift τ:

$\Delta p_1 = p(t+\tau) - p(t)$, and $\Delta p_2 = p(t+2\tau) - p(t+\tau)$.

Then we, obviously, have: $\Delta p_1 + \Delta p_2 = p(t+2\tau) - p(t)$.

Let us consider the following Variance Ratio test:

$$VR(q) = \frac{Var(q)}{q \cdot Var(1)}.$$



1) In the case $q = 2$ we have:

$\overline{(\Delta p_1 + \Delta p_2)^2} = 2 \cdot \sigma^2 \cdot (1 + \rho_1)$. Therefore, we have

$VR(2) = 1 + \rho_1$.

2) Similarly, in the case $q = 3$ we have:

$\overline{(\Delta p_1 + \Delta p_2 + \Delta p_3)^2} = 3 \cdot \sigma^2 + 4 \cdot \sigma^2 \cdot \rho_1 + 2 \cdot \sigma^2 \cdot \rho_2$, from which follows that

$$VR(3) = 1 + \frac{4}{3}\rho_1 + \frac{2}{3}\rho_2.$$

3) Further, for any integer q one can show that:

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

Therefore, using this test, we can easily distinguish between trend-following and mean-reverting statistical properties of a time series by the slope of the test $VR(q)$ near small time-shifts q.

* - with changes from "The Economics of Financial Markets" by Campbell, Lo and MacKinlay, ref. B5.

Basic Behavioral Biases and Price Predictabilities

- ***Over-Reaction or Mean-Reversion***, when agents over-react to new information by overselling on new bad information with later correction and/or over-buying on good new information with later opposite correction.
- ***Under-Reaction or Trend-Following***, when agents under-react to new information, by establishing a partial position, waiting for confirmations to their actions from other agents. Once received, they continue to increase their position in the same direction – self-reinforcement. Thus, through delayed chain reactions, the new information is gradually priced into the market.

Price-Change Sign Counting Experiments

- Data type used: 1-minute frequency, back-adjusted futures prices since inception (different for each market) until present.
- In both of these experiments the frequency (1-minute) is chosen to: be small enough in order to reveal the self-similar statistical properties within the continuous price assumption $p=p(t)$, and be large enough as compared to the so-called "bid-ask bounce" ("fake" mean-reversion). This can be easily verified by comparing the standard deviation of 1-minute price changes with the average ask-bid spread, the standard deviation has to be several times (5-10) larger.
- Both experiments are inspired by some of the early experiments of Andrew W. Lo.

Experiment 1: Counting Continuations and Reversals

For any particular time - separation τ, starting from the minimal of 1 minute to the maximal of 1,000 minutes, we will calculate :

$N_+(\tau)$ – the number of pairs of consequitive price changes in the same direction (Continuations);

$N_-(\tau)$ – the number of pairs of consequitive price changes in the opposite direction (Reversals).

We will also calculate the number of pairs of consecutive price changes where at least one price change is equal to zero, $N_0(\tau)$.

For any separation τ we can also measure :

p - probability to observe a positive move;

q - probability to observe a negative move; then

$(1 - p - q)$ will be the probability to observe an unchanged price move.

Experiment 1: Counting Continuations and Reversals

For non - symmetric RW and any time separation τ
the following simple formulas hold :

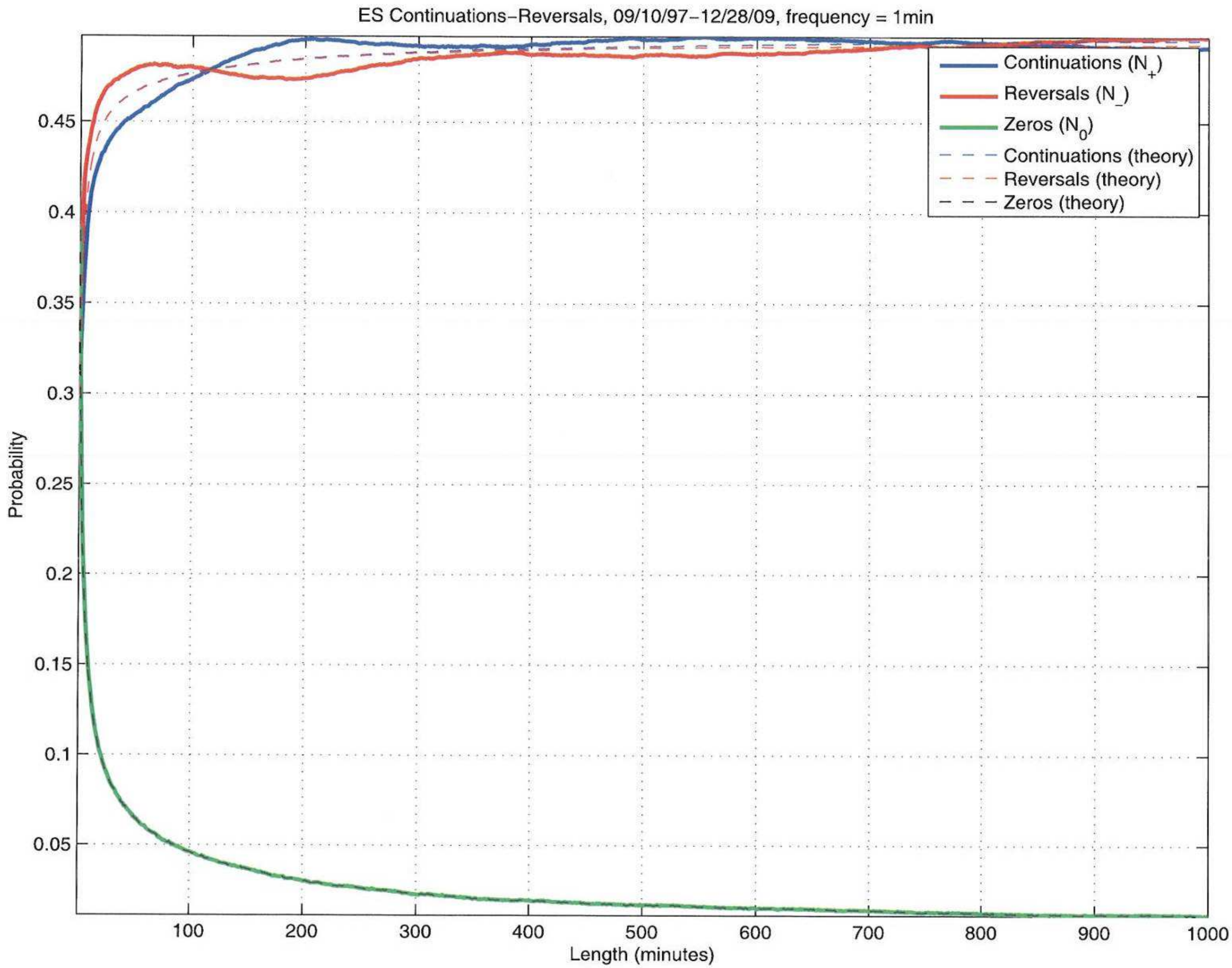
1st Price Change	2nd Price Change	Probability	Type
+	+	p^2	C
+	-	p*q	R
-	+	q*p	R
-	-	q^2	C
+	0	p*(1-p-q)	0
0	+	(1-p-q)*p	0
-	0	q*(1-p-q)	0
0	-	(1-p-q)*q	0
0	0	(1-p-q)^2	0

$$P(C) = p^2 + q^2;$$

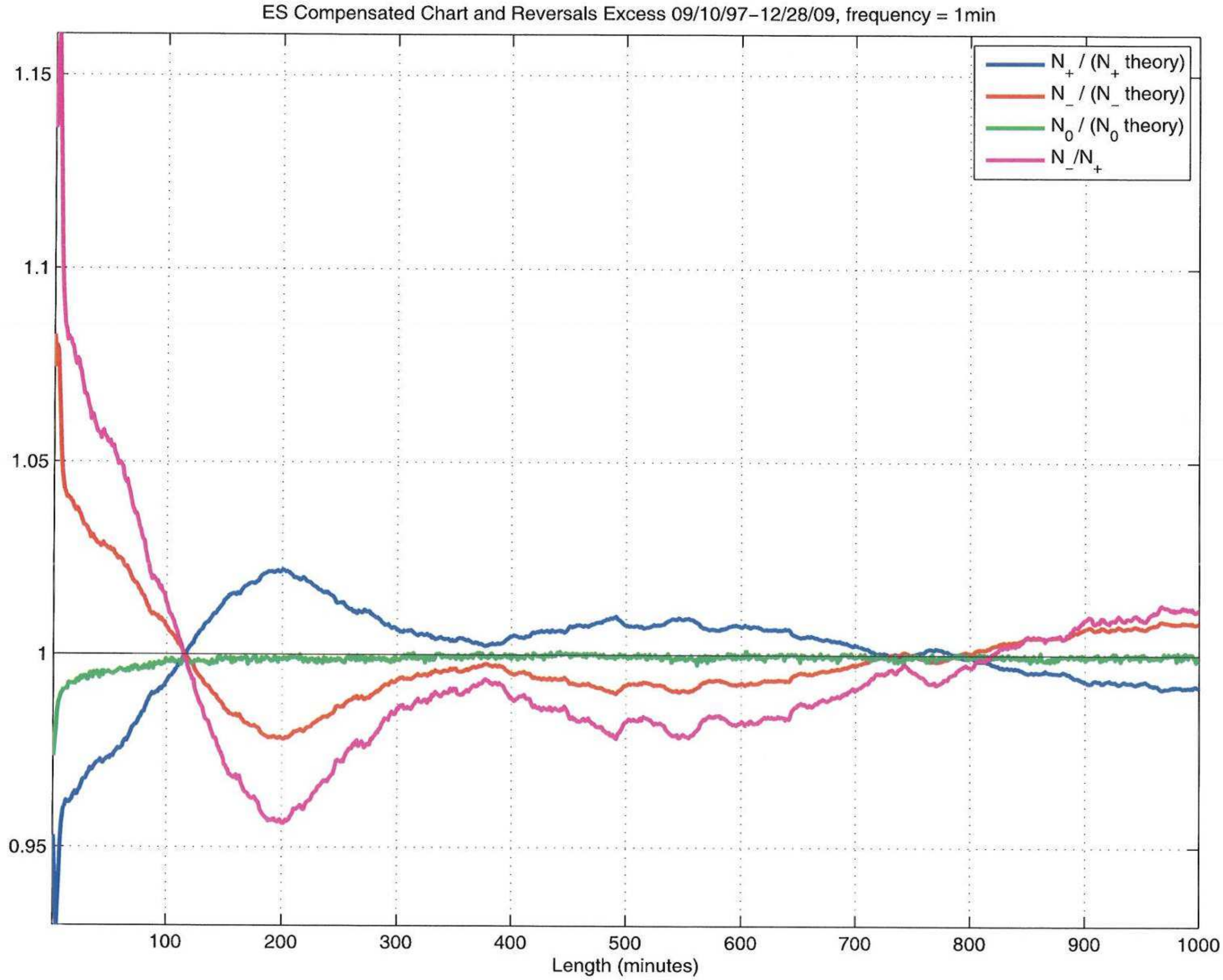
$$P(R) = 2pq;$$

$$P(0) = 1 - (p+q)^2.$$

Experiment 1: Counting Continuations and Reversals



Experiment 1: Counting Continuations and Reversals



Experiment 1: Counting Continuations and Reversals

- Evidence of short-term (up to a couple of hours) over-reaction or mean-reversion and longer-term (beyond a day) under-reaction or trend-following;
- Short-term over-reaction or mean-reversion is quite strong and robust statistically.
- Longer-term under-reaction or trend-following is weaker and less robust statistically.
- The agreement with the RW model gets better as time-separation gets larger.

Experiment 2: Counting Up- and Down- Trends

Here for any particular time - separation τ in the set $\{1\text{ min}, 5\text{min}, 15\text{min}, 1\text{hr}\}$, we will calculate :

$N_+(\tau, l)$ – the number of chains of consequitive price changes of length l

in the positive direction (up trends);

$N_-(\tau, l)$ – the number of chains of consequitive price changes of length l

in the negative direction (down trends).

We will also calculate the number of chains of consecutive zero price changes, $N_0(\tau)$.

For any separation τ we, again, will also measure :

p - probability to observe a positive move;

q - probability to observe a negative move; then

$(1 - p - q)$ will be the probability to observe an unchanged price move.

Experiment 2: Counting Up- and Down- Trends

Here, for time series length M, time separation τ, and the chain length l

are the exact solutions for those numbers of chains for an asymmetric RW :

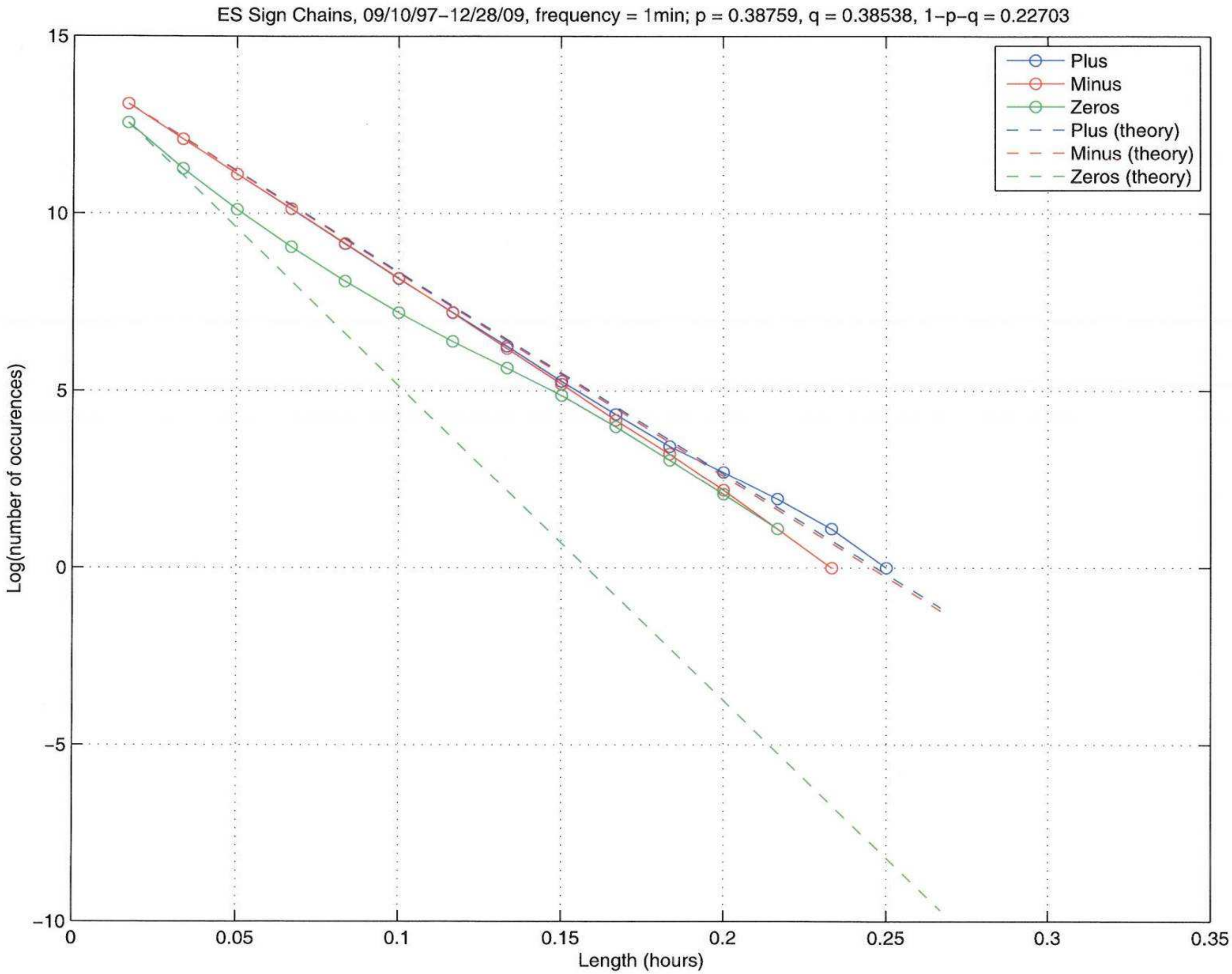
$N_{+}(\tau, L) = (\mathrm{M} - \mathrm{L} - 1) \cdot \mathrm{p}^{\mathrm{L}}$ - for the number of up - trends of length L;

$N_{-}(\tau, L) = (\mathrm{M} - \mathrm{L} - 1) \cdot \mathrm{q}^{\mathrm{L}}$ - for the number of down - trends of length L; and

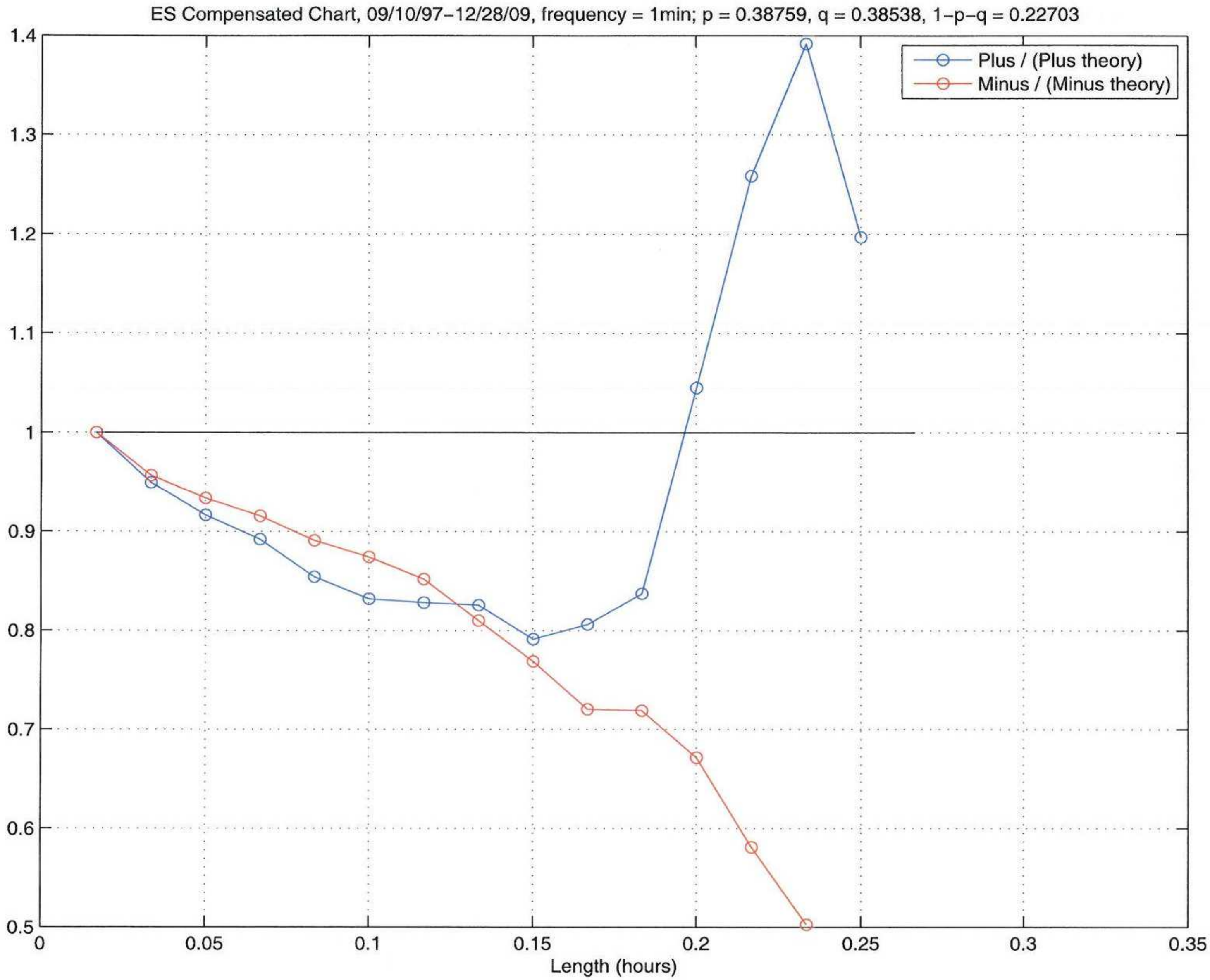
$N_{0}(\tau, l) = (\mathrm{M} - \mathrm{L} - 1) \cdot (1 - \mathrm{p} - \mathrm{q})^{\mathrm{L}}$ - for the number of flat periods of length L.

(similar distributions are called Poisson or Exponential distributions.)

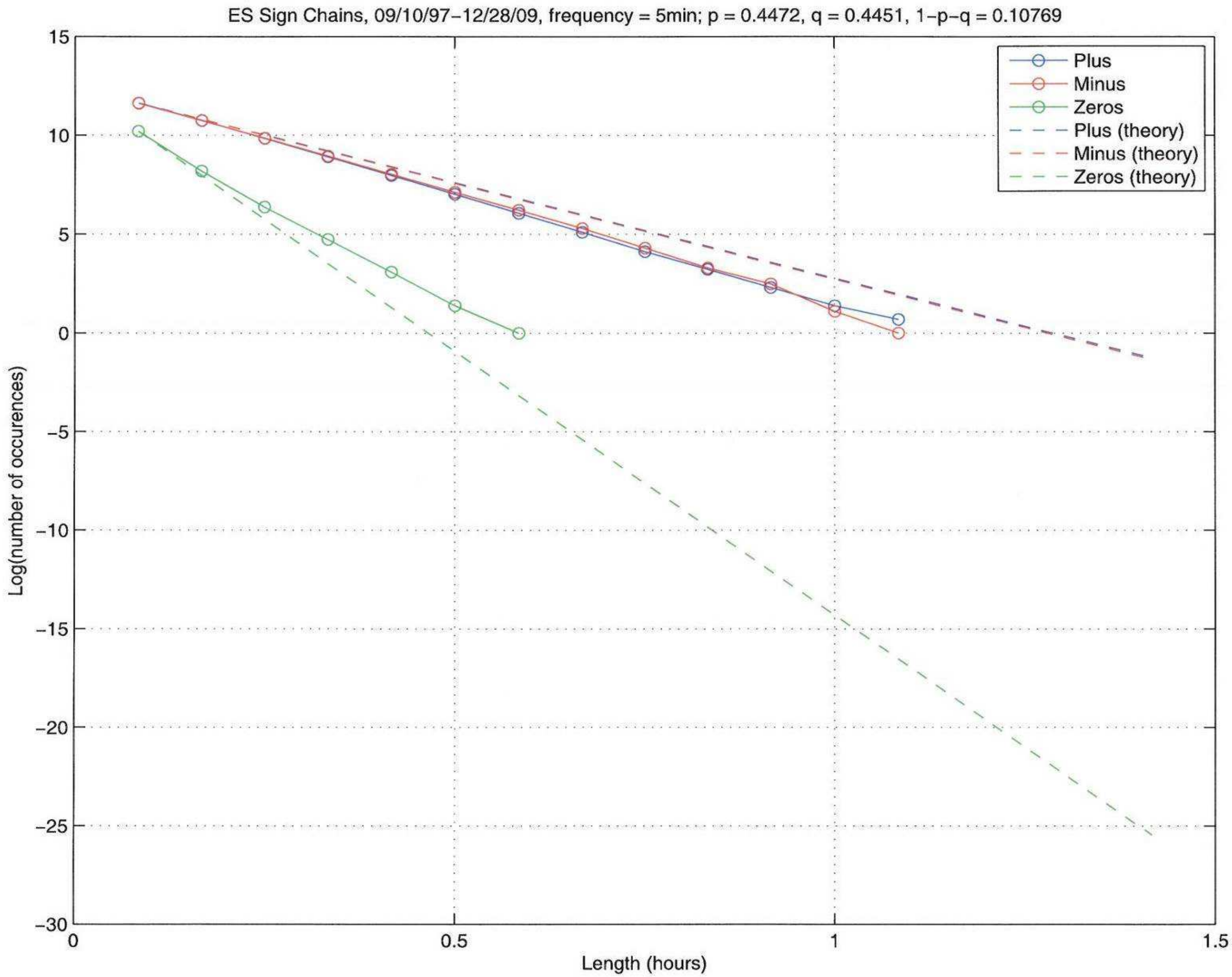
Experiment 2: 1-min



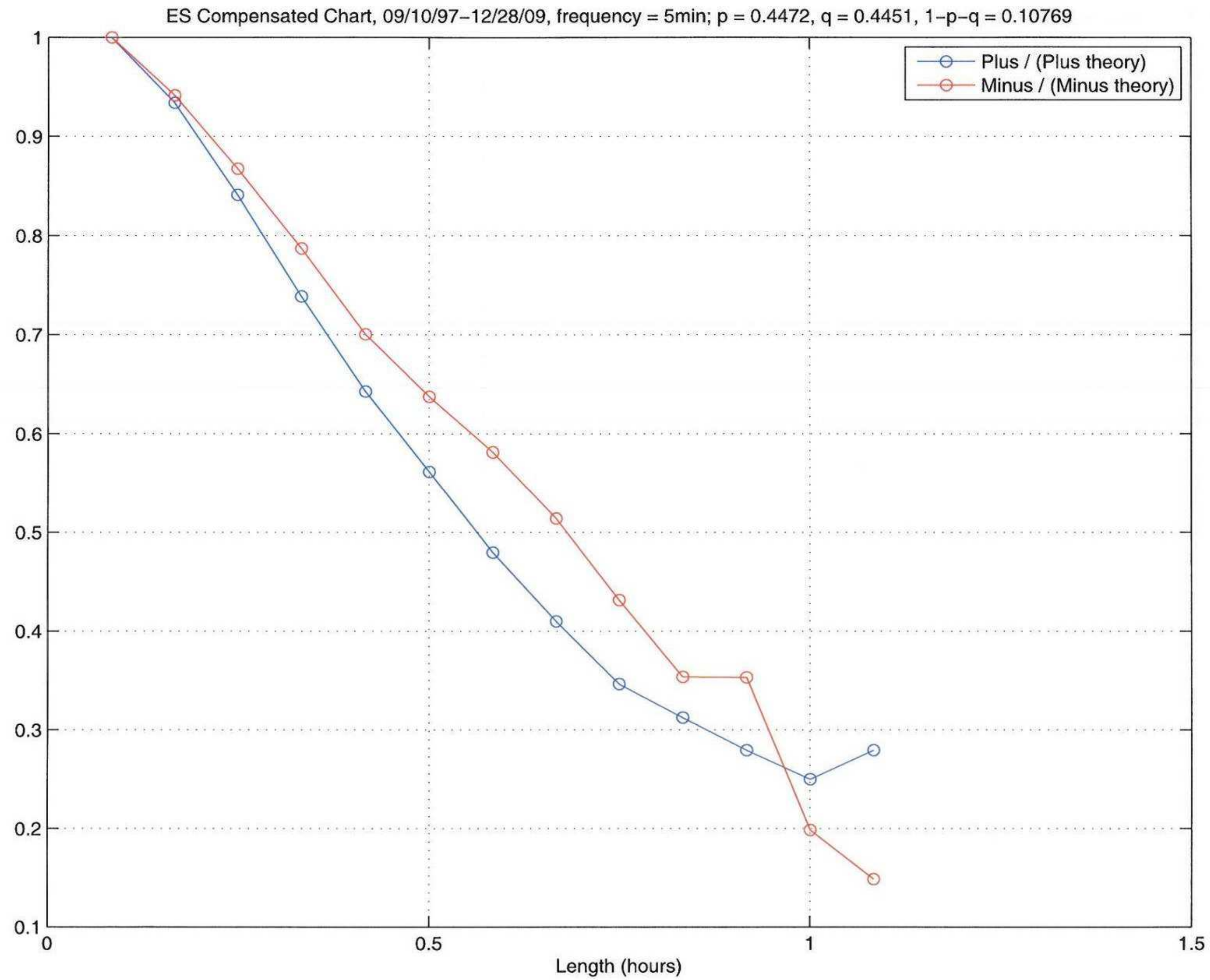
Experiment 2: 1-min



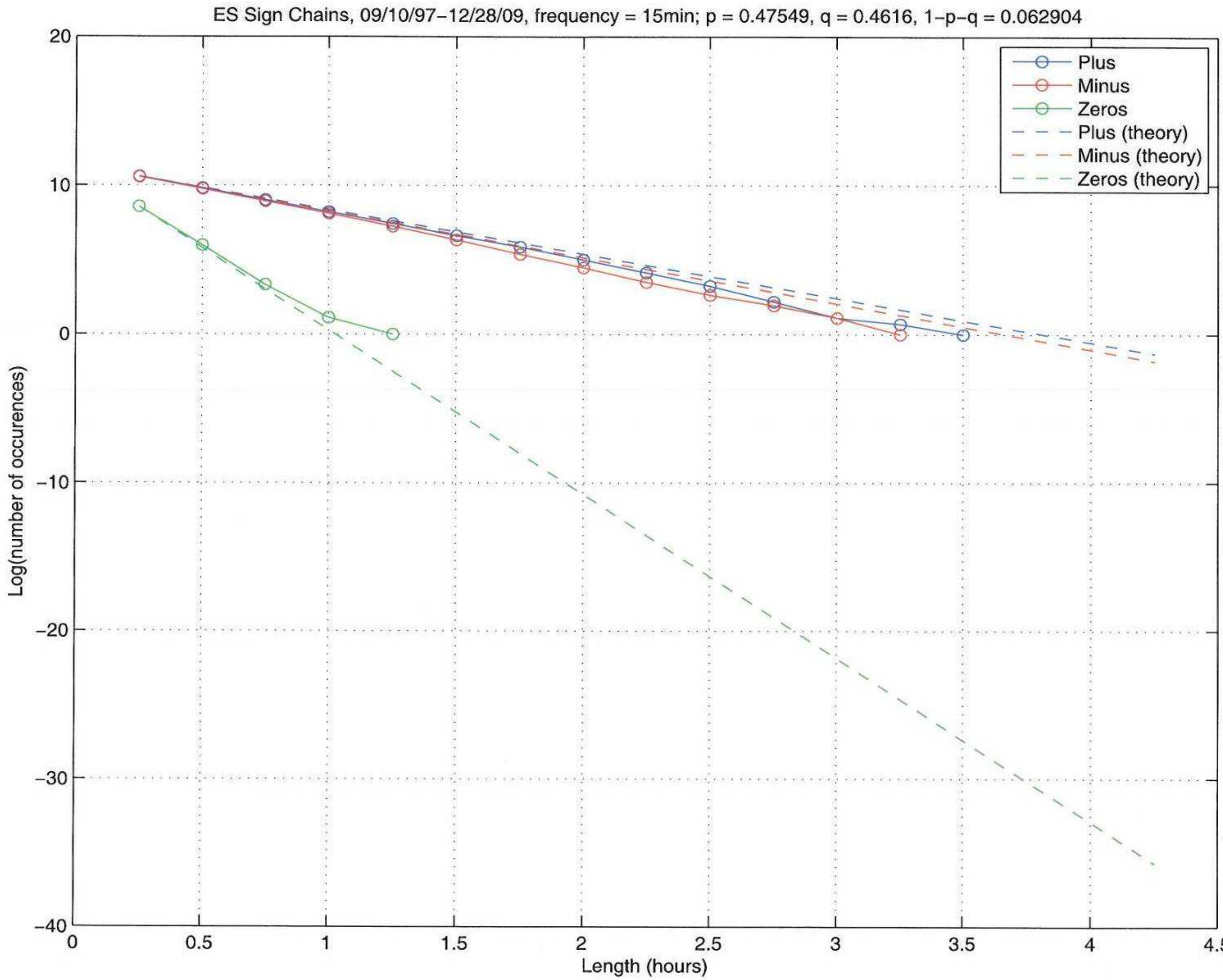
Experiment 2: 5-min



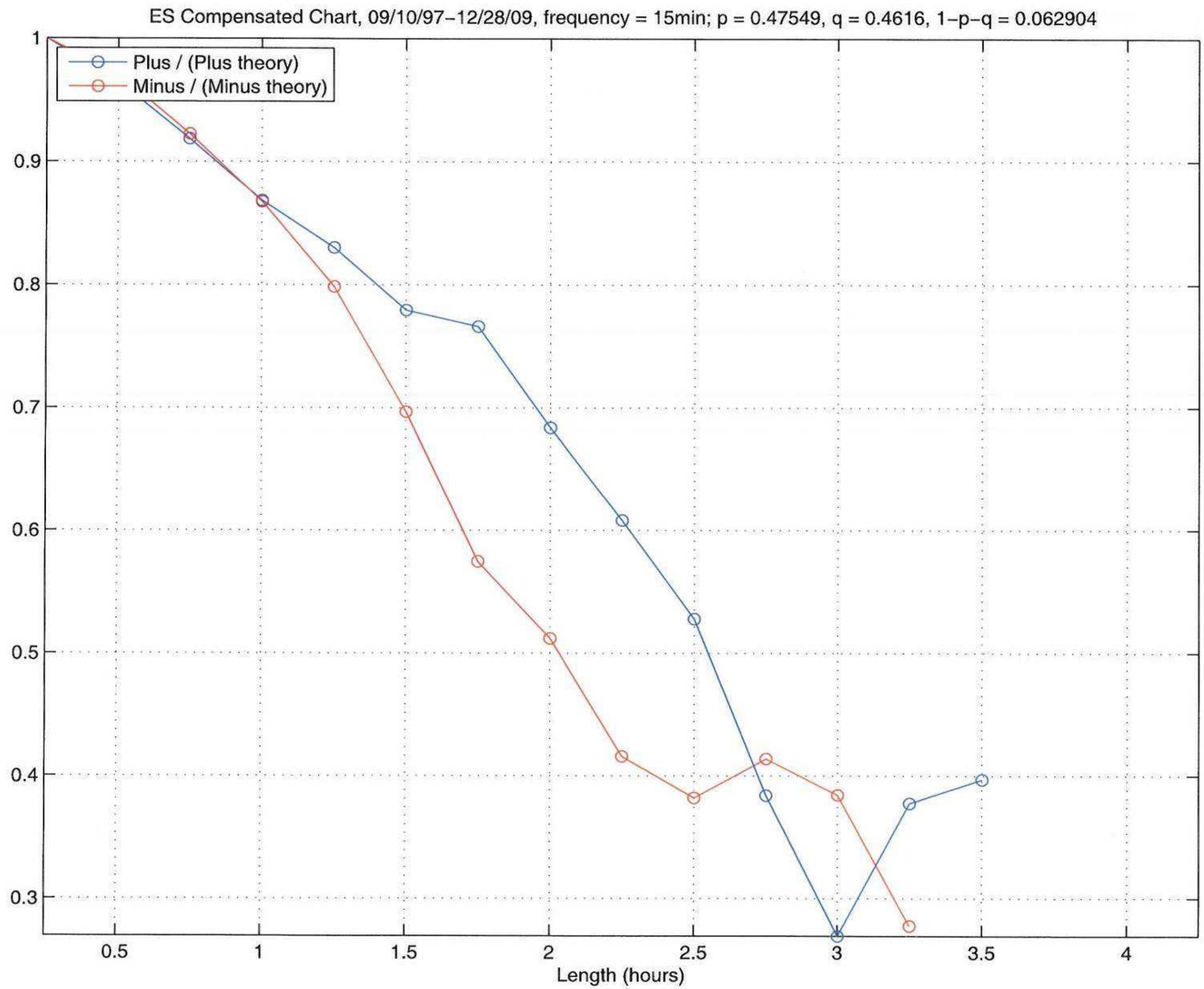
Experiment 2: 5-min



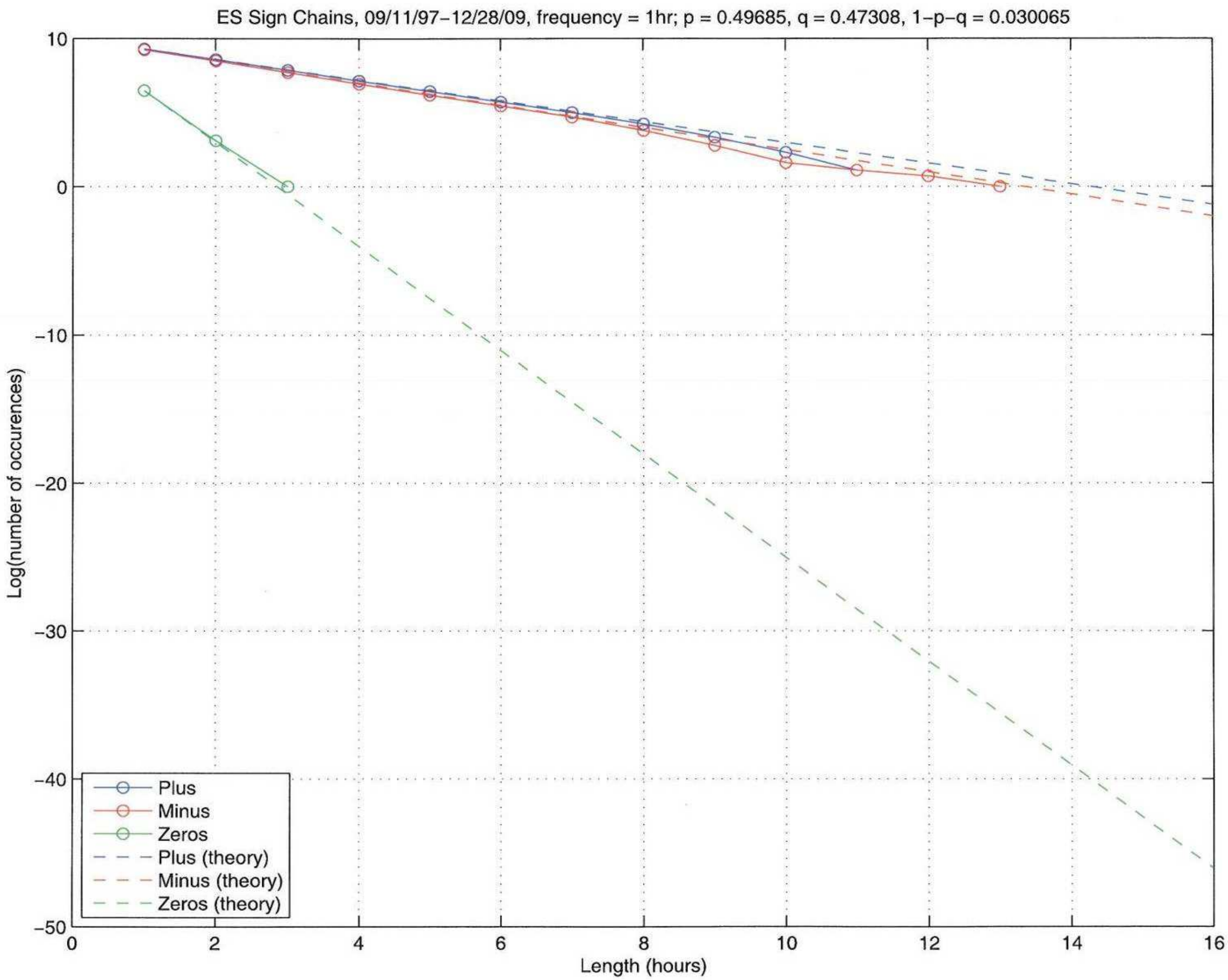
Experiment 2: 15-min



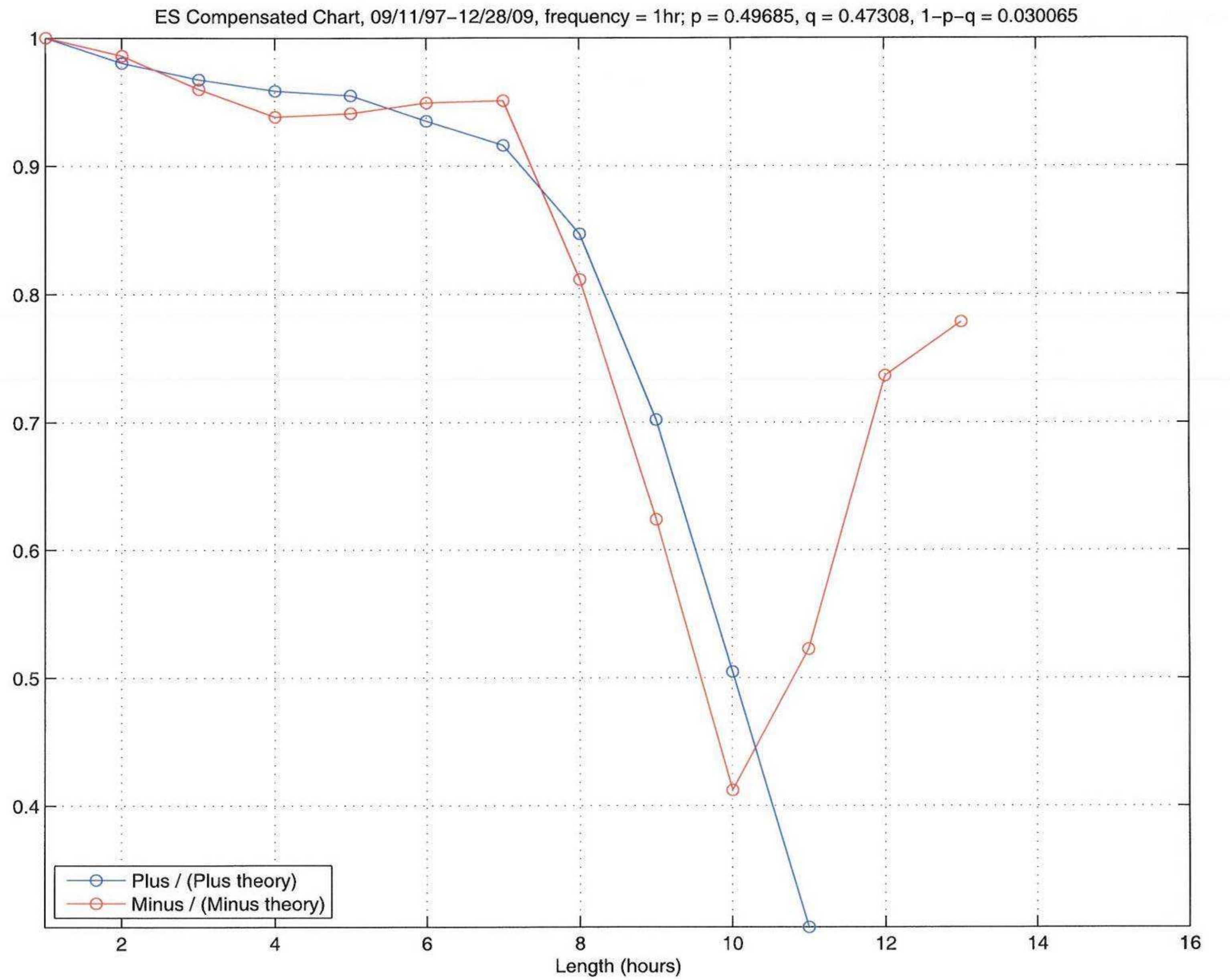
Experiment 2: 15-min



Experiment 2: 1-hour



Experiment 2: 1-hour



Experiment 2: Counting Up- and Down- Trends

- Not only the numbers of trends above theoretical values indicate under-reaction or trend-following behavior, but conversely, the numbers of trends below theoretical values indicate possible over-reaction or mean-reversion;
- There is a reasonable agreement between the two experiments, although this experiment provides further evidence on how weak the under-reaction or trend-following regime is;
- Short length trends counts have better agreement with the RW formulas.

Variance Ratio Test

- We will now transition from the signs under- and over-reaction to the price change under- and over-reaction studies;
- Time series will be taken since inception to present at 1-min resolution with time-separation from 1 min to 90 trading hours, during most liquid session (pit session);

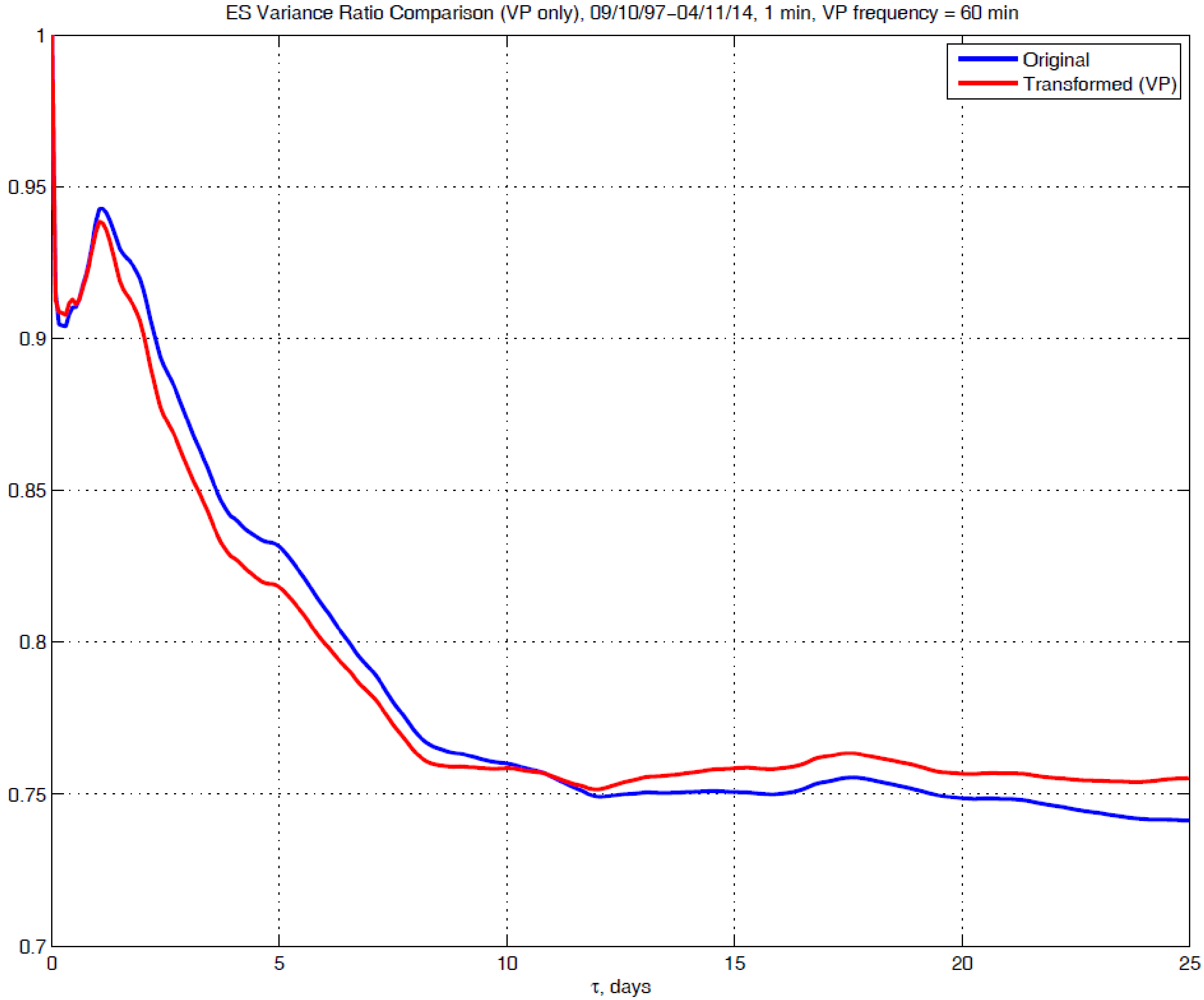
As defined in Campbell, Lo, MacKinlay, variance ratio is:

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

where q is a discrete time separation in minutes, and ρ_{k} is auto-correlation coefficient of two price changes separated by k minutes.

Time separation q is here measured in trading hours, which allows us to put them all on the same x-axys.

Variance Ratio Test



Push-Response Diagram Test

- This test is free from the fat-tailed bias of the VR test - positive;
- This test is quickly growing sample error as you increase the Δp – negative.

As defined in papers by V. Trainin et. al, "response" y is defined as :

$$\langle y \rangle_x = \int_{-\infty}^{+\infty} y \cdot P(y|x) \cdot dy,$$

as the conditional mean response to a "push" x.

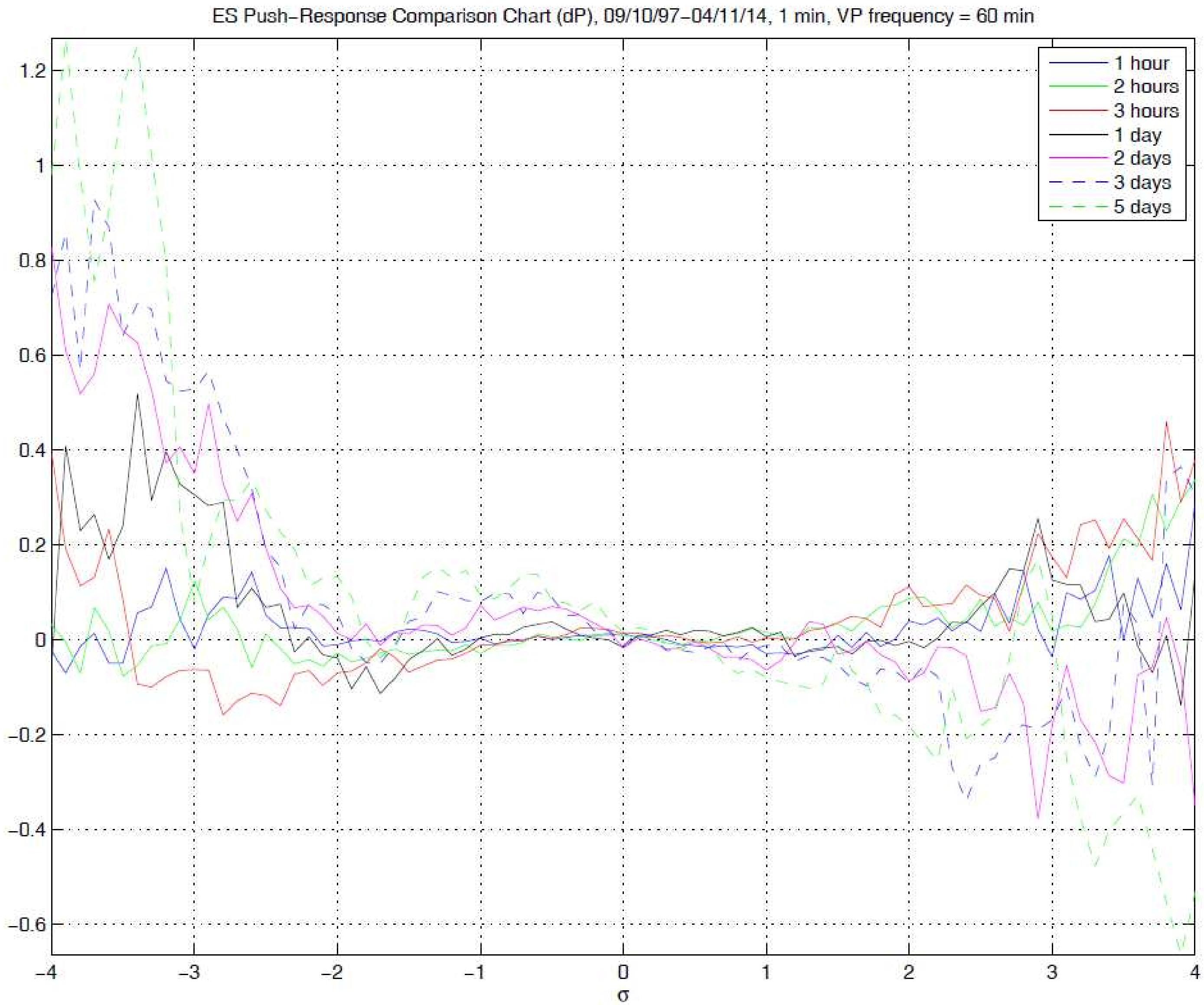
Here $P(y|x) = \dfrac{P(x, y)}{P(x)}$ is a probability density function of price change of y subject to the immediately preceding price change of x.

It can be decomposed into a symmetric and asymmetric parts, of which only the asymmetric part influences the response :

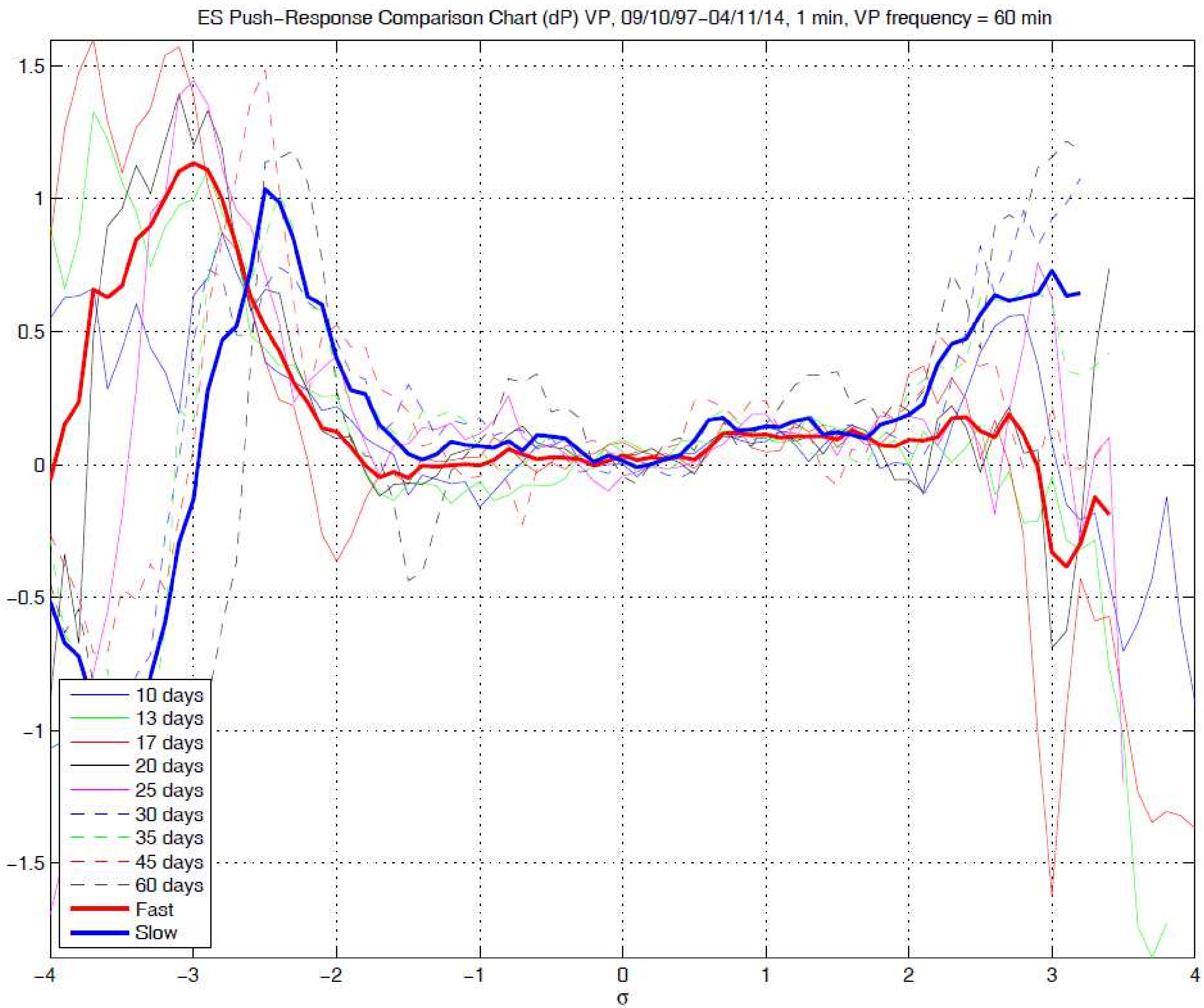
$$P(x, y) = P^s(x, y) + P^a(x, y):$$

$$P^s(x, y) = \frac{P(x, y) + P(x, -y)}{2}, \text{ and } P^a(x, y) = \frac{P(x, y) - P(x, -y)}{2}.$$

Push-Response Diagram Test



Push-Response Diagram Test



Random Walk Comparisons Tests Results

- General inspection of the test results confirms the previous sign-tests results: a general pattern is statistically strong short-term over-reaction or mean-reversion, beyond which either inconclusive or statistically weaker, selective longer-term under-reaction or trend-following properties;
- Beyond 10 trading days time-separation shows little predictability;
- These tests are more general than the first two signs tests because they considers both the price change sign and its magnitude - positive;
- These tests could be somewhat biased if the price-difference distributions function is "fat tailed" or the data sample is not large enough – negative.