ON THE NATURE OF THE STOCK MARKET: SIMULATIONS AND EXPERIMENTS

by

Hendrik J. Blok

B.Sc., University of British Columbia, 1993
M.Sc., University of British Columbia, 1995

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES

(Department of Physics and Astronomy)

We accept this dissertation as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 2000

Chapter 5

Analysis and Results: Empirical results

In the last chapter we explored the phase space of the Centralized and Decentralized Stock Exchange Models (CSEM and DSEM, respectively). This chapter is concerned with contrasting the data from these models with empirically known qualities of real markets. Some of the properties we hope to uncover are leptokurtosis in the price returns and correlated volatilities. As we will see, the emergence of these properties is closely related to the phase transitions discovered in the last chapter.

5.1 Price fluctuations

We begin by exploring the distribution of price fluctuations.

5.1.1 Background

It has long been known that stocks exhibit stochastic fluctuations in their price histories. Originally it was hypothesized that the markets exhibited (discrete) Brownian motion and therefore had Gaussian-distributed price increments [64]. Later this was adapted to account for the strictly positive nature of stock prices via geometric Brownian motion [46] with the logarithm of the price following a random walk. Much theoretical work on derivative pricing assumes geometrical Brownian motion including the famous Black-Scholes equation [65].

It was a startling discovery, then, when Mandelbrot pointed out that, empirically, the logarithm of price-returns did not have a Gaussian distribution [5, 9]. In fact, on short timescales, large (exceeding a few standard deviations) fluctuations occurred much too frequently to be explained by the Gaussian hypothesis. These

large fluctuations contribute to the tails of the distribution resulting in "fat tails".

Mandelbrot proposed the correct probability distribution function (on the logarithmic scale) was not the Gaussian but its generalization—the stable Lévy distribution. Lévy distributions (see Section C.2.3) drop off as power laws

$$p(x) \sim \frac{1}{x^{\alpha+1}},\ |x| \to \infty \tag{5.1}$$

for $0 < \alpha < 2$, resulting in fatter tails than the Gaussian ($\alpha = 2$). An attractive feature of this hypothesis is that it scales: that is, the distribution (and the exponent α) remains the same whether measured hourly, daily, or even monthly. Mandelbrot measured the daily and monthly distribution of returns from cotton prices and found both fitted well to a Lévy distribution with exponent $\alpha = 1.7$ [9]. Studies of other markets have had similar results concluding $1.4 \leq \alpha \leq 1.7$ [4, 10, and references therein]. (Appendix B demonstrates it is possible to simulate fat tails by regular sampling of a discrete Brownian process but Palágyi and Mantegna [66] demonstrate this is not responsible for the fat tails observed in return distributions.)

Since then the adequacy of the Lévy distribution to describe price fluctuations has been called into question because it implies that the fluctuations have an infinite variance whereas experimental evidence indicates it is probably finite [7,67]. (Recent studies indicate the tails of the return distributions fall off fast enough that the variance is finite.) Related to this is the observation that scaling is violated on long timescales (of more than a week [4]) where the distribution converges to a Gaussian (because the Central Limit Theorem applies if the variance is finite).

This discrepancy was initially resolved by arbitrarily truncating the power law with an exponential weighting function for large events [4, 68]. According to this theory, *gradually truncated Lévy flight* (GTLF), the tails of the cumulative distribution of (normalized) returns r follow

$$C(r) \sim \begin{cases} r^{-\alpha} & |r| \leq l_C \\ r^{-\alpha} \exp\left[-\left(\frac{|r|-l_c}{k}\right)^{\beta}\right] & |r| > l_C. \end{cases} \tag{5.2}$$

with an exponential decay beyond some cut-off l_C. While this did improve the quality of the fit to observed returns it did so at the expense of three new fitting parameters—the cut-off l_C, the decay rate k, and the power of the exponential β—bringing the total to five adjustable parameters.

An appealing alternative is the idea that the tails of the return distribution do have a power law but with an exponent $\alpha \approx 3$. (An interesting, and testable, consequence is that moments higher than three—such as the kurtosis—are divergent.) Since the exponent is greater than two the distribution is not stable and converges

to a Gaussian on long timescales. It also implies that the variance is finite. A very comprehensive analysis was performed across 1,000 companies yielding a (huge!) dataset of 40 million returns (Mandelbrot had only available 2,000 points) and the results strongly support the *inverse cubic* (IC) hypothesis [6, 69].

To be precise, the theory is that a Lévy law ($\alpha \approx 1.4$) applies for small to intermediate returns (less than a few standard deviations) but then the distribution crosses over to the inverse cubic for larger returns. Thus we have two fitting parameters in either scaling regime and a crossover point, for a total of five adjustable parameters, the same as GTLF. Although this research [6,7,69] is excellent the practical application of the theory is somewhat cumbersome and a simpler alternative exists.

5.1.2 Alternative: Decaying power law

In this section an alternative which appears to explain the data almost as well but with only three parameters is presented. Koponen [70]—expanding on work done by Mantegna and Stanley [71]—demonstrated that a power law with a smooth exponential cutoff has Lévy increments on short timescales which converge to Gaussian after a long time. This process is equivalent to GTLF with $l_C = 0$ and $\beta = 1$ so there exists no cut-off point, the exponential truncation applies for all returns. The hypothesis is that the tails of the cumulative return distribution obey

$$C(|r_i| > r) \sim r^{-\alpha} e^{-r/r_c} \tag{5.3}$$

where α is the scaling exponent and r_c is the decay constant, which sets a characteristic scale over which the power law dominates—in the limit $r_c \to \infty$ Mandelbrot's pure Lévy flight hypothesis re-emerges. (This functional form has been observed to accurately describe the distribution of fluctuations in the "game of Life" [72].)

The use of this truncation hypothesis must be justified in light of the overwhelming empirical evidence supporting the inverse cubic hypothesis [6,7,69]. To do so it is necessary to explicitly formulate the goals of this section: (1) to determine if the return distributions generated by the models scale with an exponent near $\alpha \approx 1.4$ over some range of returns, and, if so, (2) to estimate the range of returns over which the scaling holds.

In doing so we can determine if the models reproduces truncated Lévy flight observed empirically. However, the amount of data collected will be insufficient to adequately determine how the truncation occurs, whether it is an inverse cubic or exponential decay. Further, this detail is not of central importance to this work so it is left for future research.

Therefore we are free to choose the most convenient form for the truncation factor, that given by Eq. 5.3. This form has a few technical advantages over the previously discussed methods. The first is that it requires a fit of only three parameters and the fit is linear in each of them when performed on the logarithmic scale. Hence, only one optimal solution exists and a number of algorithms exist for arriving there [20, Ch. 15].

Secondly, it is a single continuous function so it requires no manual searching for a crossover point between two regimes (which is an art in itself). In fact, it automatically determines the crossover from power law behaviour to exponential decay with the fitted parameter r_c. The larger r_c is, the greater range the power law is valid over.

Fig. 5.1 contrasts the fit of the inverse cubic hypothesis (a) with the decaying power law (b). Both fit the high frequency exchange data quite well, but recall the inverse cubic requires two additional parameters to do so. The decaying power law hypothesis indicates a power law with exponent $\alpha \approx 1.42$ applies for returns less than $r_c \approx 8.11$ (standard deviations), beyond which the power law is attenuated by an exponential decay.

Also shown is the distribution of daily returns for the Nasdaq Composite index over almost 16 years [73, ticker symbol=^IXIC] in Fig. 5.2. This figure demonstrates an important point: when the crossover to the exponential occurs at a small value of r_c (less than a few standard deviations, as in the positive tail) the estimate of the Lévy exponent is unreliable. The larger r_c gets, the more meaningful the value of α becomes.

While the claim that real market fluctuations actually obey Eq. 5.3 is largely unsubstantiated, a weaker claim that this functional form is an effective method to test for scaling in market data is also being made based on two observations: (1) it is systematic and does not require any intervention (tuning of parameters) on the part of the researcher, and (2) it characterizes both the range and exponent of the Lévy region well with the parameters r_c and α, respectively. For these reasons this method will be used in the following sections to test for scaling in CSEM and DSEM.

5.1.3 Methodology

To determine the distribution of returns over some time interval Δt the price series will be regularly sampled producing a series of returns

$$r_i \equiv \ln\left[\frac{p((i+1)\Delta t)}{p(i\Delta t)}\right]. \tag{5.4}$$

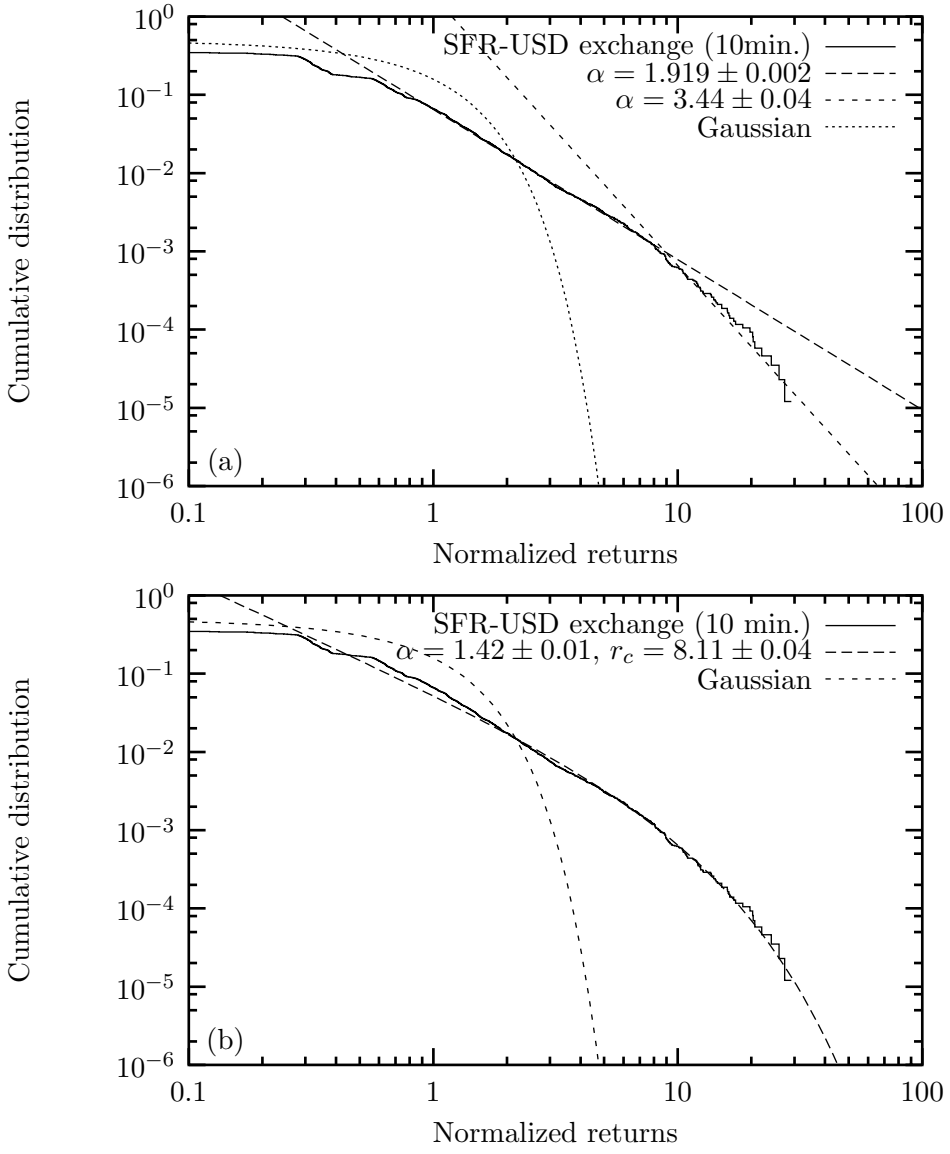

Figure 5.1: Ten minute returns (86,000 data points) of the Swiss franc–U.S. dollar exchange rate [2] (negative tail) compared to power law with crossover to $\alpha \approx 3$ (a) and power law with exponential drop-off presented in this section (b).

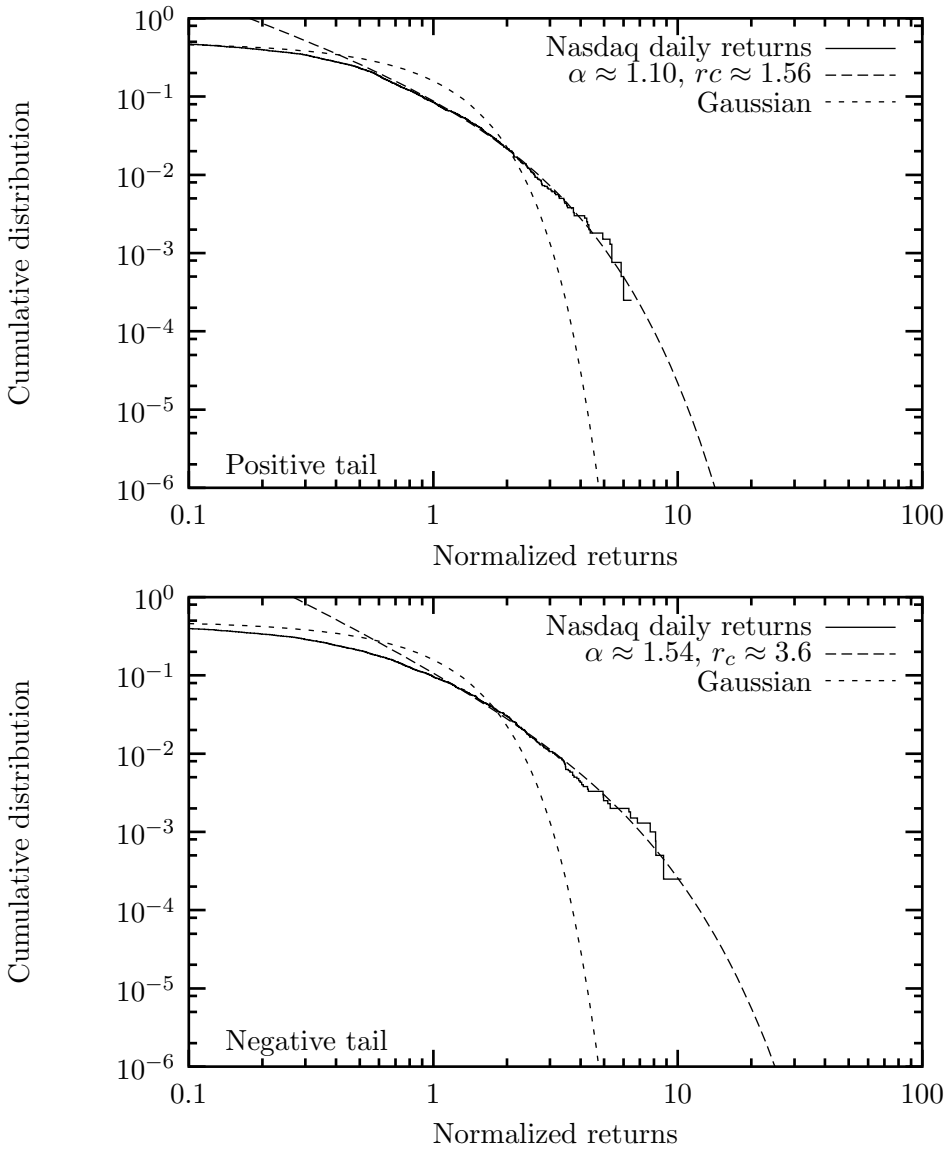

Figure 5.2: Both tails of the cumulative distribution of daily (normalized) returns for the Nasdaq Composite index between October 1984 and Jun 2000 (4,000 data points) fit well to a decaying power law. The power law is truncated by two standard deviations in the positive tail but extends almost to four in the negative tail.

The mean $\bar{r}$ and standard deviation σ_r will be computed and the returns normalized

$$\hat{r}_i \equiv \frac{r_i - \bar{r}}{\sigma_r}. \tag{5.5}$$

The cumulative distribution of the normalized returns will be calculated and compared with the cumulative Gaussian (the error function). Of particular interest are the tails of the distribution which are hypothesized to obey the scaling functions

$$C(\hat{r}_i \geq \hat{r}) \sim \quad \hat{r}^{-\alpha_+} \exp(-\hat{r}/r_{c,+}), \qquad \hat{r} \to +\infty \tag{5.6}$$

$$C(\hat{r}_i \leq \hat{r}) \sim \quad |\hat{r}|^{-\alpha_-} \exp(-|\hat{r}|/r_{c,-}), \quad \hat{r} \to -\infty \tag{5.7}$$

with exponents α_+ and α_- for the positive and negative tails and crossover values $r_{c,+}$ and $r_{c,-}$. (The cumulative distribution is preferred because cumulating effectively "smooths" the data, making it more amenable to analysis.)

The adjustable parameters $\alpha_\pm$ and $r_{c,\pm}$ will be acquired via a Levenburg-Marquardt nonlinear fit to $\log C$ for returns exceeding $|\hat{r}| > 1$ since we only want to fit the tails. (A linear fit is also possible with a suitable choice of parameters.)

A small value of r_c will be interpreted to mean that no scaling exists and the parameter α is irrelevant.

5.1.4 Centralized stock exchange model

For this experiment we return to Dataset 1 (Table 4.1) and apply the analysis to the largest system $N = 1000$. Fortunately, each run consists of over 30,000 days worth of data so the scaling can be tested on a wide range of timescales. (In the analysis, the initial transient will be discarded.)

Since CSEM contains the free parameter σ_ϵ (the forecast error) we must sample a suitable spectrum of values in our search for scaling. Obviously, the dynamics around the critical point $\sigma_c \approx 0.08$ (from Fig. 4.6(a)) is of particular interest so samples are chosen which span the critical point.

The scaling regime is indicated by the characteristic return r_c: a small value indicates that the exponential drop-off occurs for small returns (before scaling becomes evident) and a large value indicates the power law applies over a broad range of returns. In this experiment a threshold value of $r_c = 3$ was observed to adequately distinguish between distributions which scaled and those which didn't. This limit was also used in Ref. [7] to estimate the scaling exponent α.

Plotting the characteristic return r_c for a variety of forecast errors σ_ϵ (Fig. 5.3) shows that scaling is only observed well below the critical point σ_c. The average scaling exponent for all distributions with $r_c > 3$ is $\alpha = 0.8 \pm 0.4$, which compares poorly with the empirical value $1.4 \leq \alpha \leq 1.7$.

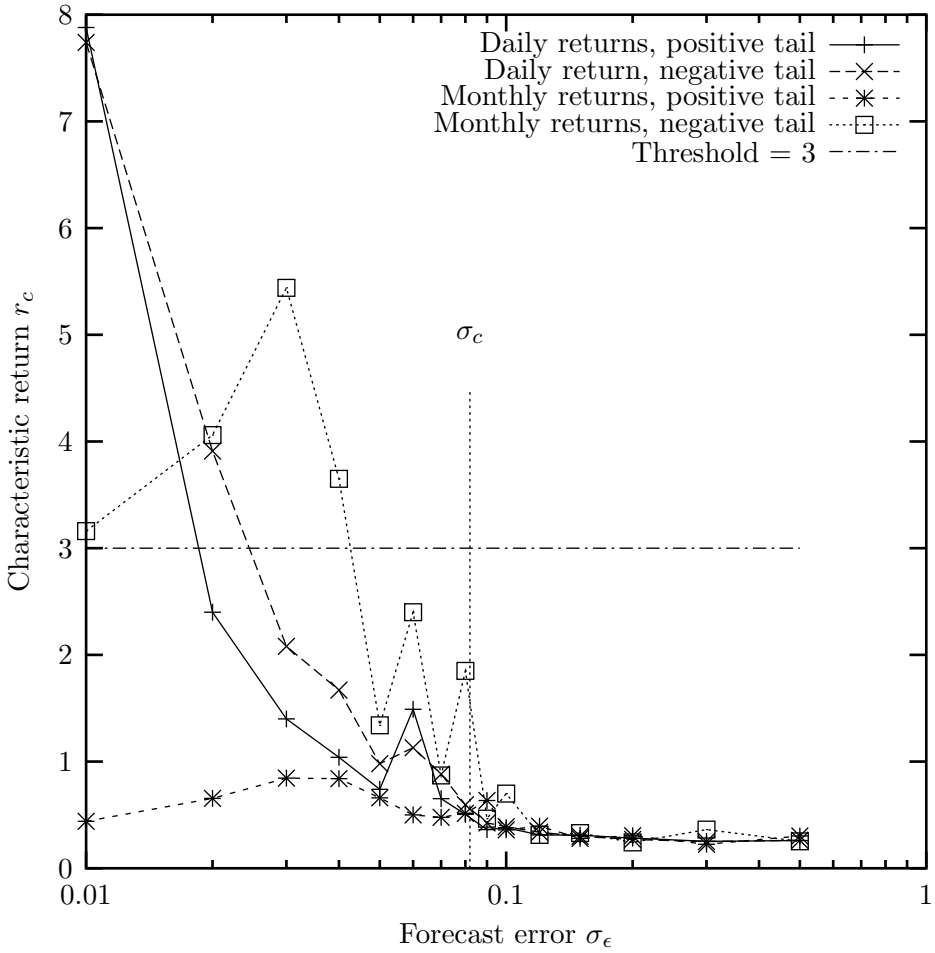

Figure 5.3: Scaling in the distribution of returns is only observed well below the critical point $\sigma_\epsilon \ll \sigma_c$ in CSEM as indicated by large values of the characteristic return r_c. For small σ_ϵ scaling occurs in both tails for daily returns but only for negative returns in monthly returns.

Parameters	DSEM Dataset 3		
Particular values			
News response r_n	0.01	0.01	0.001
Price response r_p	-0.75 to 0.75 by 0.25	0.95	0.99
Number of runs	7	1	1
Common values			
Number of agents N	100		
Run length ("days")	20,000		

Table 5.1: Parameter values for DSEM Dataset 3. These runs are a variation of Dataset 1 (all unspecified parameters are duplicated from Table 4.4) run out for longer times (roughly 80 years). Also notice that for $r_p = 0.99$ the news response was reduced by an order of magnitude to keep the price within reason.

For positive returns scaling is found to disappear as the sampling interval is increased from daily to monthly (20 days), as expected. Interestingly, the same is not true for negative returns: instead the returns scale for even more values of σ_ϵ on a monthly timescale than they do daily.

The run $\sigma_\epsilon = 0.03$ sampled monthly is an interesting case because it exhibits a strong asymmetry between up and down moves, having a characteristic return above the threshold for negative returns and below for positive returns, so its return distribution is plotted in Fig. 5.4. This effect is due to an asymmetry between up- and down- movements which arises from the artificial price cap imposed by the parameter δ. See, for example, Fig. 4.1(b) which shows that occasional large crashes occur when the price approaches its upper limit while upwards movements are more normally distributed.

Unfortunately, the above only serves to further call in to question the validity of the CSEM model because scaling behaviour is only observed below the critical point, in the regime where we have already seen (Fig. 4.1, for instance) the dynamics are completely unrealistic. So we turn to DSEM in the hope that it is a more realistic model of market dynamics.

5.1.5 Decentralized stock exchange model

In this section the distribution of returns in DSEM is analyzed.

To get the limit distribution a large quantity of data is required so DSEM was run with the parameter values from Table 5.1, the most notable feature being that the run length was extended from 1,000 days ($\approx$ 4 years) to 20,000 days ($\approx$ 80

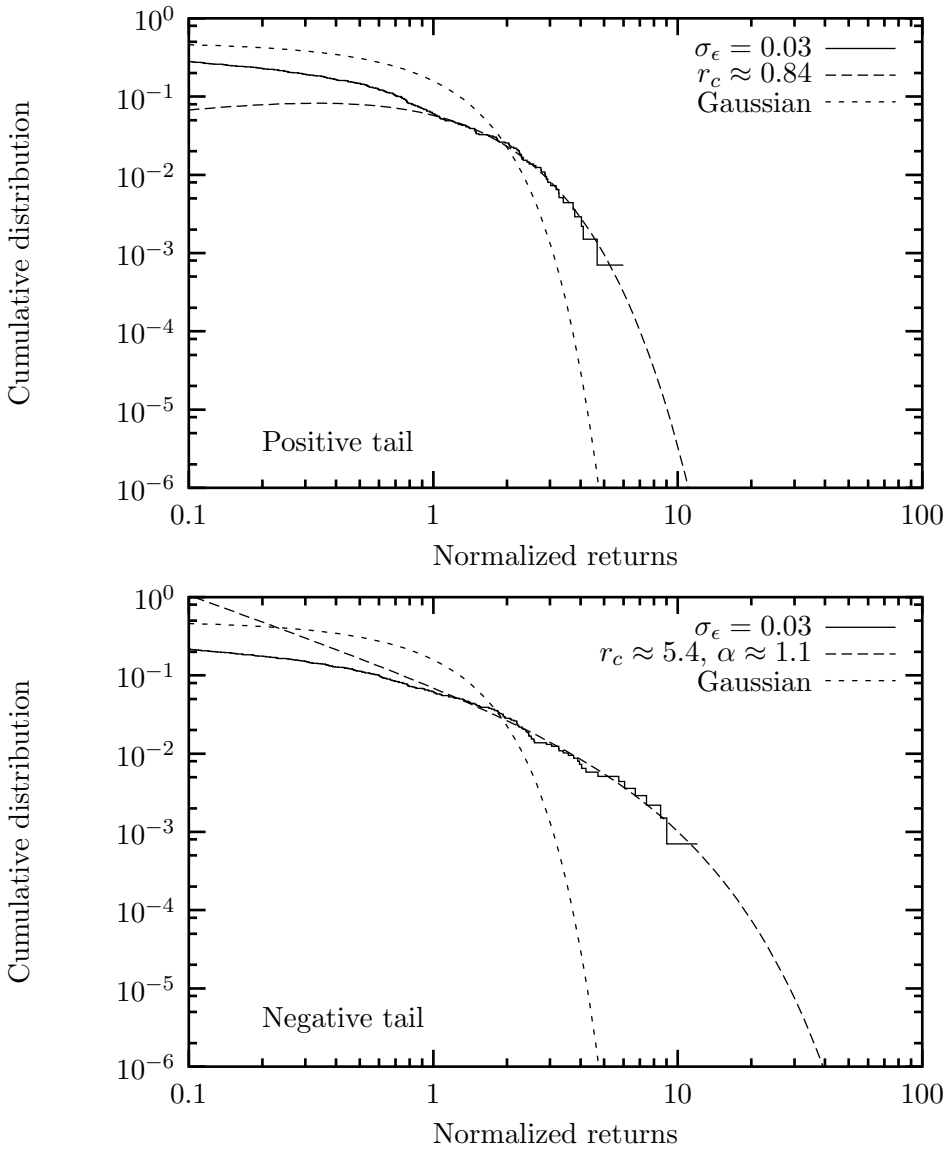

Figure 5.4: For $\sigma_\epsilon = 0.03$ in CSEM ($N = 1000$) the distribution of positive (monthly) returns (upper) almost converges to a Gaussian but still has a slightly heavy tail. The negative returns (lower), however, exhibit scaling for $r < r_c \approx 5.4$ with an exponent $\alpha \approx 1.1$.

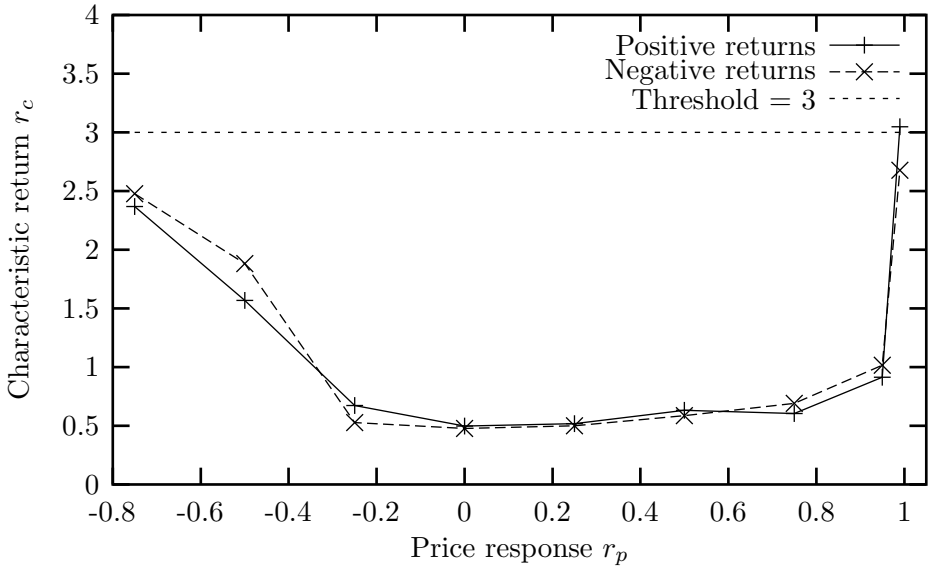

Figure 5.5: DSEM only begins to exhibit scaling, as measured by a characteristic return exceeding three standard deviations, for price responses well below the first-order transition $r_2 = -0.33$ and as the price response approaches the critical point $r_1 = 1$.

years).

The distribution of price returns (sampled "daily") was then cumulated and fitted with Eq. 5.3. (Issues raised in Appendix B regarding sampling are addressed on page 122, below.) The characteristic size of the returns for the different values of r_p is shown in Fig. 5.5. Notice that they are almost exclusively below the threshold required to establish scaling indicating that the distributions do not exhibit scaling properties observed empirically. The worst region appears to be intermediate values of r_p with better performance near the endpoints.

Recall that DSEM exhibits three behaviours as r_p is varied: (1) when $r_p > r_1 = 1$ the price is perfectly autocorrelated—every movement is followed by another (typically larger) movement in the same direction; (2) in the intermediate region $r_1 > r_p > r_2$ the price series looks most realistic and has (at most) weak autocorrelations on long timescales; and (3) when $r_p < r_2 \approx -0.33$ the price fluctuations have a strong negative autocorrelation extending over all timescales. Thus, the price fluctuations appear only to obey (realistic) scaling distributions in the domains precisely where the dynamics were observed to be unrealistic! To reconcile this dichotomy we need to expand our experimental parameter space.

Parameters	DSEM Dataset 4
Number of agents N	100
Lower price response r_{lo}	–0.75 to 0.75 by 0.25
Higher price response r_{hi}	0.50 to 1.50 by 0.25
Number of runs	35
Run length ("days")	20,000

Table 5.2: Parameter values for DSEM Dataset 4. These runs are characterized by a two-point distribution of the price response. Each agent chooses $r_p = r_{lo}$ or r_{hi} with equal probability. (All unspecified parameters are duplicated from Table 4.4.)

Two-point price response

Thus far the price response had been fixed at a single value for all the agents. But since realistic dynamics (characterized by both the lack of strong memory effects and scaling in the distribution of returns) were not to be obtained by any single value of r_p I was forced to allow multiple price responses. Originally I explored allowing r_p to span a broad range which covered all three phases but the range required to get scaling was so large that most of the agents were either in Phase 1 ($r_p > r_1$) or in Phase 3 ($r_p < r_2$) with only a few in Phase 2. Therefore it seemed easier to just require that r_p take on one of only two allowed values, r_{lo} and r_{hi}.

Data analysis

Thirty five runs were executed spanning a two-dimensional region of parameter space with each agent choosing a price response of either r_{lo} or r_{hi} (with equal probability). The lower price response was varied between $-0.75 \leq r_{lo} \leq 0.75$, spanning the first order phase transition at $r_2 \approx -0.33$, and the upper value was varied between $0.50 \leq r_{hi} \leq 1.50$, spanning the critical point at $r_1 = 1$, as indicated in Table 5.2.

For each run the cumulative distributions of returns (both positive and negative) were calculated and the tails (returns exceeding one standard deviation) fitted to a decaying power law (Eq. 5.3). As before, the tail was determined to scale if the decay constant r_c exceeded three standard deviations, otherwise the region over which a power law is suitable is insubstantial.

For returns smaller than the characteristic return r_c the exponential in Eq. 5.3 is almost flat so the power law dominates. Larger values of r_c indicate that scaling spans a greater range of returns. As shown in Fig. 5.6(a), scaling is observed for some parameter combinations. To test which parameter produces scaling a linear

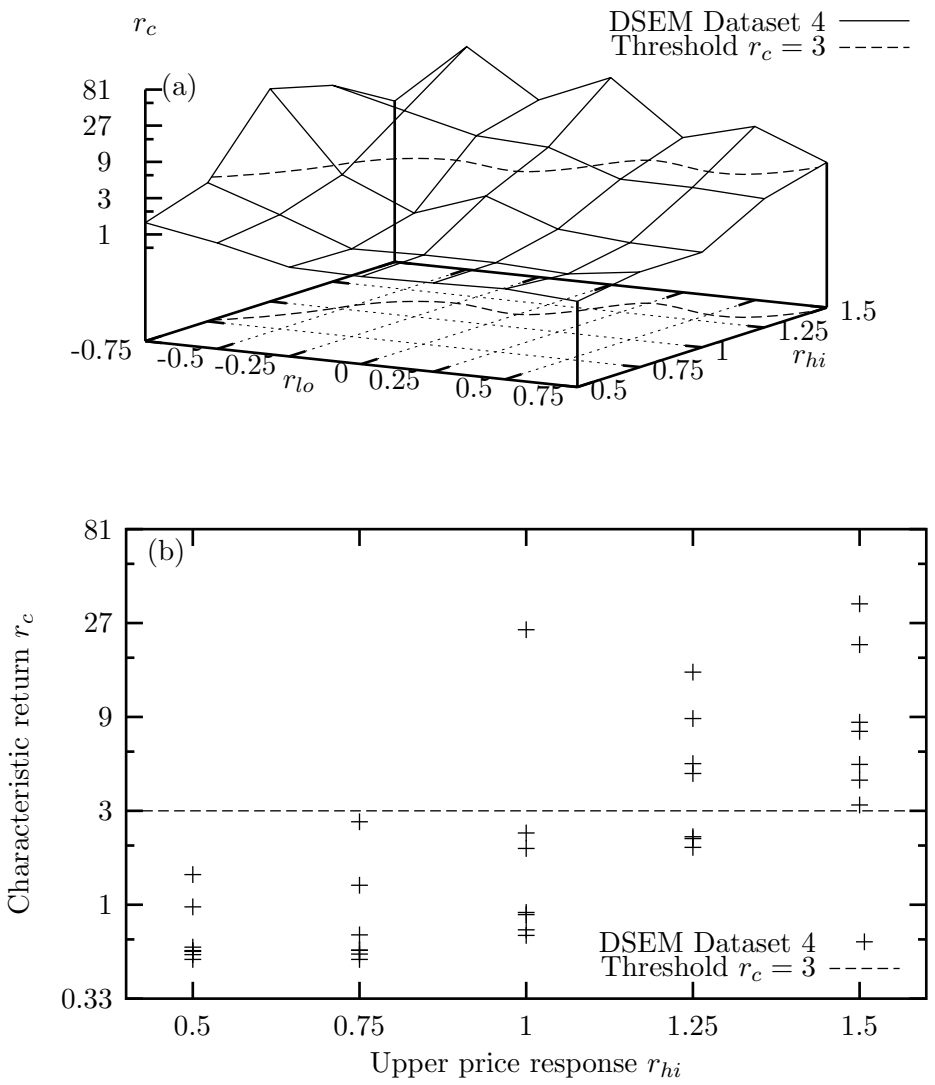

Figure 5.6: The characteristic returns in DSEM with a two-point distribution of price responses (r_{lo} and r_{hi}) exceeds the required threshold of $r_c = 3$ when r_{hi} is large (a). Neglecting the dependence on r_{lo} (b) it becomes clear that the characteristic return grows exponentially with the upper limit r_{hi}, crossing the threshold near $r_{hi} \approx 1$.

Variable	Correlation with $\log r_c$
r_{hi}	78%
$r_{hi} - r_{lo}$	76%
r_{lo}	–41%

Table 5.3: Linear correlation analysis between said variable and the logarithm of the characteristic return from DSEM Dataset 4. The correlation is strongest with the upper limit of the price response r_{hi}.

correlation analysis was performed, as shown in Table 5.3, between the logarithm of the characteristic return and a few obvious possibilities: the upper price response r_{hi}, the lower limit r_{lo}, and the spread $r_{hi} - r_{lo}$. The best predictor for scaling over a large range of returns was found to be r_{hi} with a correlation of 78%. (The spread also correlated well but, as will be seen later in this chapter, is unable to account for other empirical qualities of the market.)

Fig. 5.6(b) shows the dependence of the characteristic return on the upper price response. Notice that 85% of the data points lie in the upper-right and lower-left quadrants if axes are drawn at $r_c = 3$ (horizontal) and $1 < r_{hi} < 1.25$ (vertical). Thus, the strongest condition for scaling appears to be that the upper price response $r_{hi} > 1$, above the critical point $r_1 = 1$.

Of all the runs which exhibit scaling the average scaling exponent was calculated to be $\alpha = 1.64 \pm 0.25$, in line with the empirical value $\alpha \approx 1.40 \pm 0.05$ [10].

Recanting continuous heterogeneity

Some other market models characterize agents by types: either *fundamentalists* or *chartists*. In the derivation of DSEM I claimed (Section 3.3.3) that allowing a continuous range of the parameter r_p would be superior, reflecting a greater diversity of opinion as would be expected in the real world. However, as we saw above, DSEM is only able to capture the essence of real market fluctuations (scaling) when r_p is set to two discrete values, rather than a continuum. (As mentioned before, a continuous range of r_p can also produce scaling but only if the spread is set to a much greater value than required by the two point distribution.) In other words, scaling appears to depend on the separation of agents into "types".

Timescales

An interesting empirical property of scaling in real markets is that the exponent appears to be invariant when measured on different timescales (except for timescales exceeding a few days when the distribution converges to a Gaussian). To test if this also occurred in DSEM the run $r_{lo} = 0.00$, $r_{hi} = 1.25$ was chosen for further analysis. This run was chosen because it was observed to exhibit scaling on timescales of one day, and because a value of $r_{lo} = 0$ seems "natural"—it divides the population into two types: one of which are pure fundamentalists, not responding to price fluctuations at all.

This run was sampled at ten different intervals ranging from 0.02 days (roughly 8 minutes, assuming a 6.5 hour trading day), to 20 days (one month, neglecting weekends). As can be seen in Fig. 5.7(a) the characteristic return increases with smaller sampling intervals, and drops below the threshold for detecting scaling when the interval exceeds 5 days (one week). On longer timescales the distribution indeed converges to a Gaussian (not shown).

As expected, (when scaling is detectable) the power law exponent does not appear to depend on the sampling interval, fluctuating around $\alpha = 1.55 \pm 0.11$.

Tickwise returns

Appendix B demonstrates that it is possible to generate the illusion of fat tails in a discrete Brownian process simply by sampling it at regular intervals, as was done for DSEM. Therefore, it is important to establish that the fat tails discussed above are not an artifact of sampling, but are inherent to the fluctuations themselves. This is easily tested by simply sampling the process in *trading time* rather than real time. That is, a sample is taken directly after every trade (or *tick*).

Clark [74] raised the issue of whether regular sampling may be producing the fat tails observed empirically but Palágyi and Mantegna [66] demonstrated fat tails are still observed when sampled in trading time. To test if this was also the case for DSEM Fig. 5.6(a) was reproduced, using trading time instead of daily samples, in Fig. 5.8. Clearly, scaling is still evident (in the same region of parameter space) when sampling in trading time so it is not an artifact of the sampling interval.

5.1.6 Summary

The distribution of price returns, measured as the logarithm of the ratio of successive prices, was the subject of investigation in this section.

Two theoretical curves meant to describe the tails of the distribution were presented. An alternate form was also presented, whose main advantages are that it

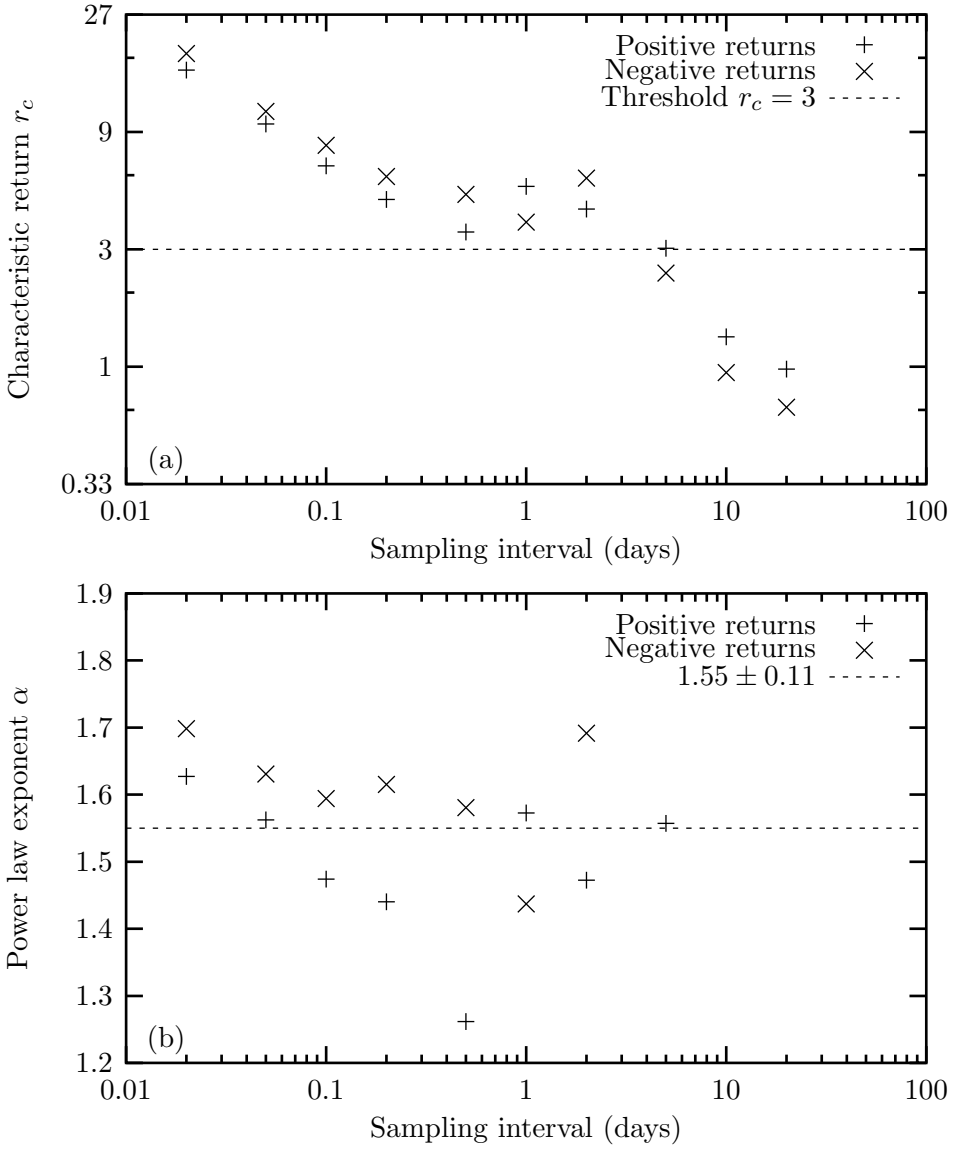

Figure 5.7: The characteristic return r_c (a) and scaling exponent α (b) for DSEM with $r_{lo} = 0.00$ and $r_{hi} = 1.25$. The characteristic return grows as the sampling interval is shortened, but the scaling exponent α is fairly constant (1.55 ± 0.11).

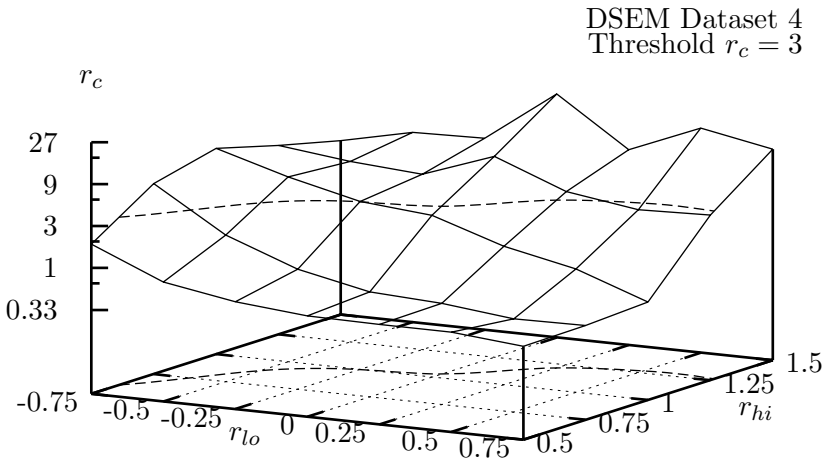

Figure 5.8: Fitting the decaying power law to DSEM with a two-point price response using returns on individual trades (rather than per unit time, as in Fig. 5.6) shows scaling still occurs in the same region of parameter space.

is linear in its parameters (of which there are two fewer than the competing models). This curve appears to describe empirical data quite well, but even if it is found to be inaccurate, it is still useful because it provides a simple, mechanical method for estimating over what range the power law dependence applies ($|r| < r_c$) and the scaling exponent itself.

Both CSEM and DSEM were tested for "fat tails" with this functional form with the requirement that the scaling extend for at least three standard deviations ($r_c \geq 3$) to be deemed significant. CSEM was found only to exhibit scaling for $\sigma_\epsilon \ll \sigma_c$, in a region of parameter space where the dynamics are known to be unrealistic.

DSEM provided some surprises: if all the agents maintained an identical price response parameter r_p then scaling did not occur except as r_p moved into regions known to produce unrealistic dynamics. However, if two values of r_p were allowed, with each agent randomly picking one or the other, scaling was observed when the responses spanned the critical point $r_p = r_1 = 1$. A test for a variety of values of r_p established a scaling exponent $\alpha = 1.64 \pm 0.25$, with returns sampled daily, comparing favourably with the empirical quantity $\alpha = 1.40 \pm 0.05$ [10]. The scaling exponent was shown to be robust, independent of the sampling interval.

In short, DSEM was able to produce realistic return distributions while

CSEM was not. Furthermore, DSEM implied the mechanism which produces scaling is somehow related to having different "types" of agents interacting with each other: fundamentalists versus chartists, for instance. The crucial determinant for scaling appeared to be that the range of parameter values spanned the critical point.

In the next section we explore a related phenomenon: autocorrelations in the price series.

5.2 Price autocorrelation

5.2.1 Background: The efficient market hypothesis

In this section we explore the possibility of serial correlations in the price series. As discussed in the last section it has long been thought that the market behaves as a random walk [46, 64]. Related to this is the *efficient market hypothesis* (EMH) which, in its weakest form, states that new information received by investors is reflected in the stock's price almost instantly [75]. Since new information cannot be predicted, neither can the future price of the stock so price movements should be independent of their histories. If this were not the case then there would be a riskless way to exploit one's foreknowledge for profit (an *arbitrage* opportunity).

The presence of transaction costs allows an even weaker form of the EMH: there may exist arbitrage opportunities (autocorrelations) but they are so small (brief) that any potential profit would be absorbed by commissions. This form of the EMH is supported by evidence: a number of studies have concluded that autocorrelations in the price series decay exponentially over a scale of only a few minutes [4, 33, 36, 76–81] (or a few trades [66]).

In this section we will look for correlations in the price series generated by DSEM. (Since it was established in the last section that CSEM is not a realistic market model an autocorrelation analysis of its price series will be dispensed with.) Of interest are both short- and long-range correlations.

5.2.2 News

Before analyzing the correlations in the price series it should be reiterated that the price series is driven by news releases such that the logarithm of the price $p(t)$ roughly follows the cumulative news $\eta(t)$ as

$$\log p(t) \propto \eta(t). \tag{5.8}$$

In Section 3.6.1 it was argued that the proportionality constant should be $r_n/(1 - r_p)$ but with a two-point price response distribution this is not applicable,

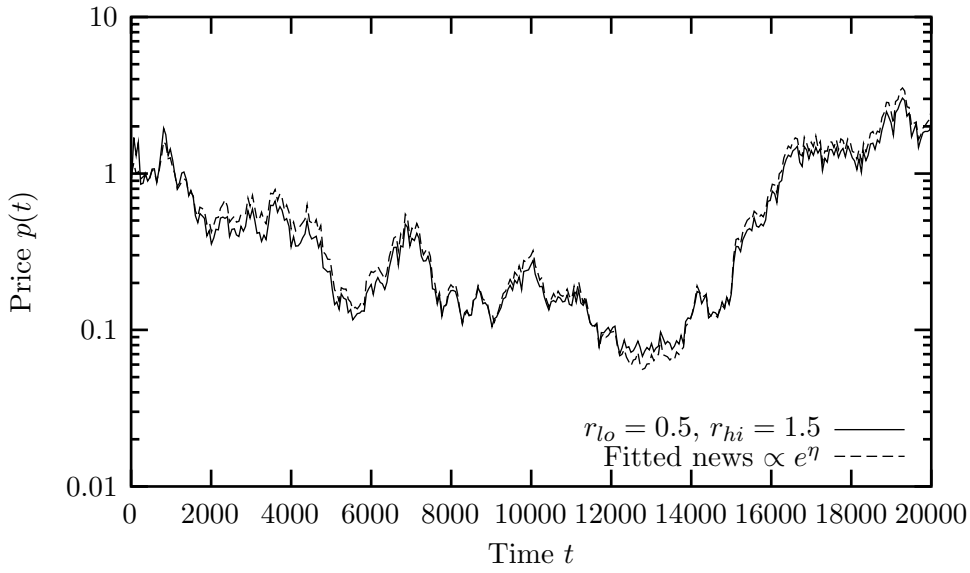

Figure 5.9: Sample price series for DSEM Dataset 4 ($r_{lo} = 0.5$, $r_{hi} = 1.5$) showing the price roughly tracks the exponential of the cumulative news e^{η}. The proportionality constant is estimated from the data.

especially given that the upper limit can be $r_{hi} \geq 1$. Nevertheless the relation still holds but the proportionality constant is best estimated from the data as is done in Fig. 5.9.

Since the news is a discrete Brownian motion we may naively expect the price series to be a simple geometric Brownian motion but as we have already seen the price series exhibits an abundance of outliers not observed in the news. As we will see in the next section, the price series also contains a memory which the news does not.

5.2.3 Short timescales

The analysis for short timescales is fairly straightforward. We need only compute the autocorrelation between returns for different lags. For this analysis, *trading time* (or *ticks*), defined as the number of transactions executed, will be used as the time index since the quantity of interest is the correlation between successive trades. For comparison, the autocorrelations between daily returns for the Dow Jones Industrial Average for the last hundred years [3] is shown in Fig. 5.10, indicating no significant correlations in support of the efficient market hypothesis. (Of course, this does not preclude correlations existing on timescales of less than one day but data were not available at this resolution.)

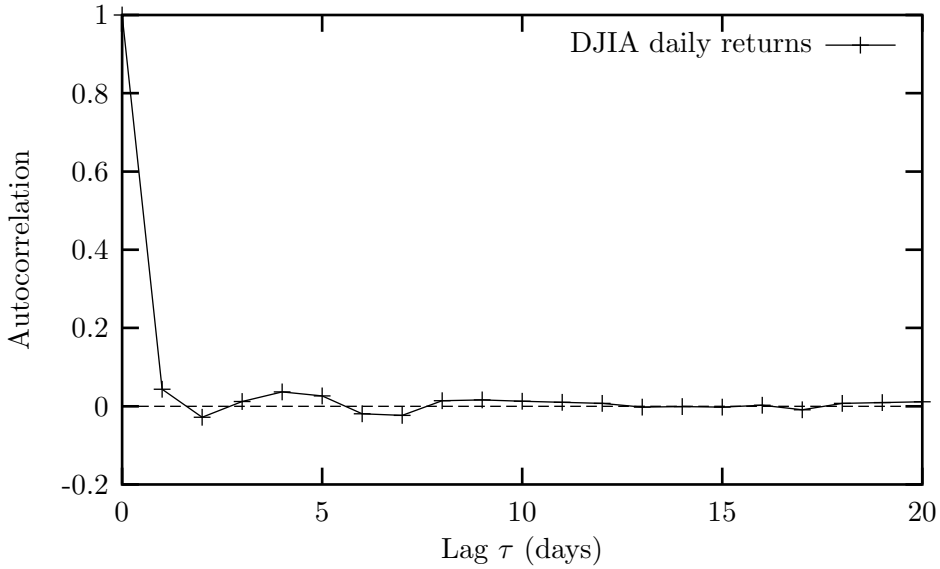

Figure 5.10: The autocorrelation between daily returns for the Dow Jones Industrial Average [3] decays rapidly to zero with an estimated characteristic timescale $\tau_c = 0.4 \pm 0.2$ days. (Being less than the sampling interval, this estimate is not precise.)

As demonstrated in Fig. 5.11 correlations also decay quickly for DSEM regardless of the value(s) of the price response, with correlations only evident over a few successive trades. This would seem to imply that the price series has no memory. However, recall that in Section 4.2.3 we observed two phase transitions in DSEM by directly measuring the memory of the price series, which challenges the results presented here.

There are two possible reasons for the discrepancy: (1) the autocorrelations are measured tickwise whereas the Hurst exponent was originally measured from a daily sample, or (2) a plot of the autocorrelation does not fully describe temporal dependencies in the data. To determine which is the case the long-range dependencies are again estimated from the Hurst exponent, this time calculated from the tickwise data.

5.2.4 Long timescales

To test for long-range temporal dependencies we again compute the Hurst exponents for the price returns in DSEM. But first it should be mentioned that data from real markets have been found to have no memory, with Hurst exponents near $H \approx 0.5$. Indeed, for the daily Dow Jones Industrial Average returns presented above, the Hurst exponent is estimated at $H = 0.484 \pm 0.013$ indicating no long-term memory

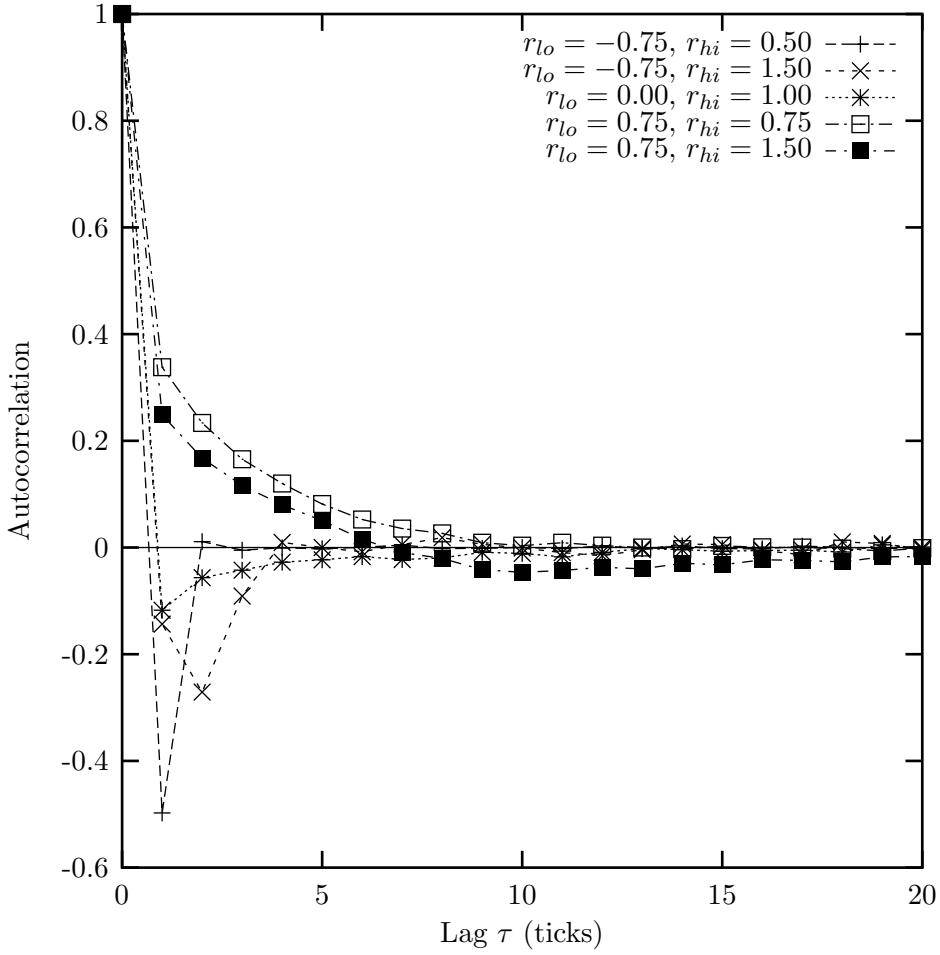

Figure 5.11: The autocorrelation between tickwise returns for DSEM (with a two-point price response distribution) decays rapidly to zero for all runs sampled.

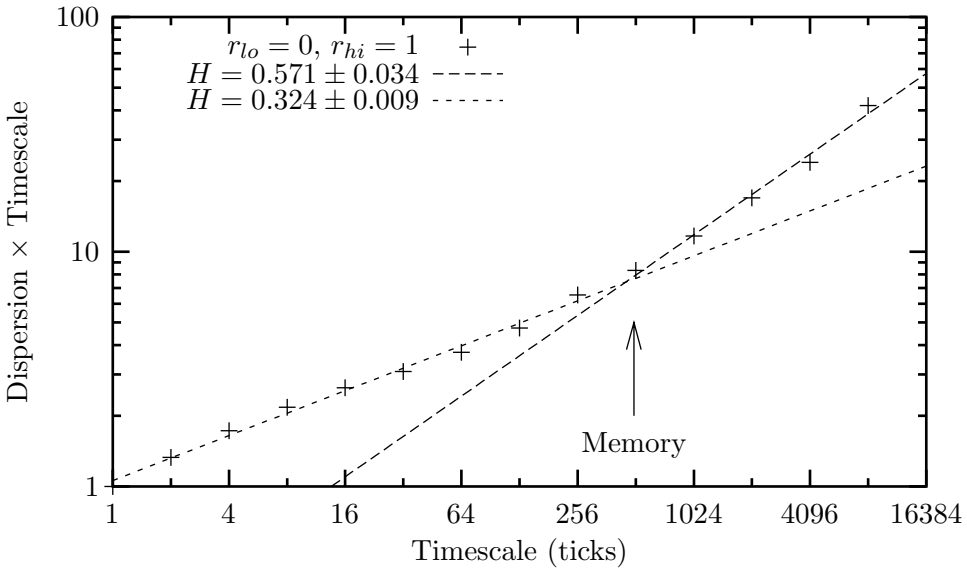

Figure 5.12: Sample dispersion plot (see Section C.2.1) demonstrating the phenomenon of crossover in the Hurst exponent to $H \approx 1/2$ on long timescales for DSEM with a two-point price response distribution.

effects.

Crossover

On first glance the computed Hurst exponents appeared similar to the original results shown in Fig. 4.11 but on closer inspection some interesting qualities were revealed. Namely, in almost all the runs there appeared to be two different scaling behaviours: for small timescales one Hurst exponent dominated but as the timescale grew there appeared a crossover to a different exponent. The latter of these was invariably near $H = 1/2$ indicating a lack of memory. A sample graph demonstrating crossover is presented in Fig. 5.12.

The reader may be concerned that the timescale used for calculating the memory is not linear but *trading time*—the cumulative number of trades executed since the start of the experiment—and the crossover phenomenon may be an artifact of this sampling. Evidence indicates that trading time is the more natural timescale [66,74], reducing biases introduced by regular sampling (see Appendix B). However, for completeness the data were also tested using regular sampling with largely the same results. A sample plot is shown in Fig. 5.13 demonstrating crossover also occurs when returns are sampled at regular intervals. The remaining discussion refers to tickwise sampling.

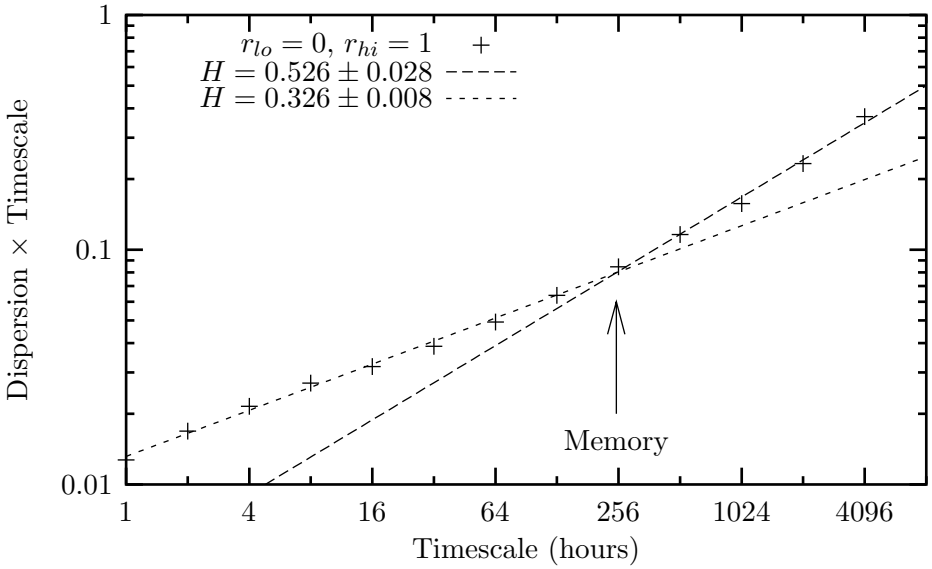

Figure 5.13: A reproduction of Fig. 5.12 except with regularly sampled returns at an "hourly" interval (instead of tickwise). Short timescale anticorrelations crossing over to uncorrelated returns at long timescales are still observed so the effect is not an artifact of sampling tickwise.

The crossover to $H \approx 1/2$ is not altogether surprising because on long enough timescales we expect the news process to be an important determination of the price movements, and the news is a simple, discretely-sampled Brownian motion with no memory ($H = 1/2$).

Yet another phase transition

The crossover point gives another estimate of the duration of correlations, or *memory* in the price series. Fig. 5.14 shows that the memory depends strongly on the lower limit of the price response r_{lo}. In fact, as this value crosses roughly $r_{lo} \approx 0.5$ a phase transition is apparent. (While interesting, this transition will not be characterized further in this thesis, but may be analyzed in future work.) For larger values of r_{lo} the memory effects disappear very quickly, conforming to the efficient market hypothesis and empirical data. Further, in this range H is already quite close to one half, which also suggests the lack of a memory for any significant period. (Personal experience suggests that the Hurst exponent has a typical error margin of ± 0.1 so any value in $0.4 \leq H \leq 0.6$ should be interpreted as potentially having no memory.)

Thus, the price returns in DSEM are observed to exhibit a realistic lack of (significant) memory when r_{lo} exceeds roughly 0.25. If DSEM is interpreted as

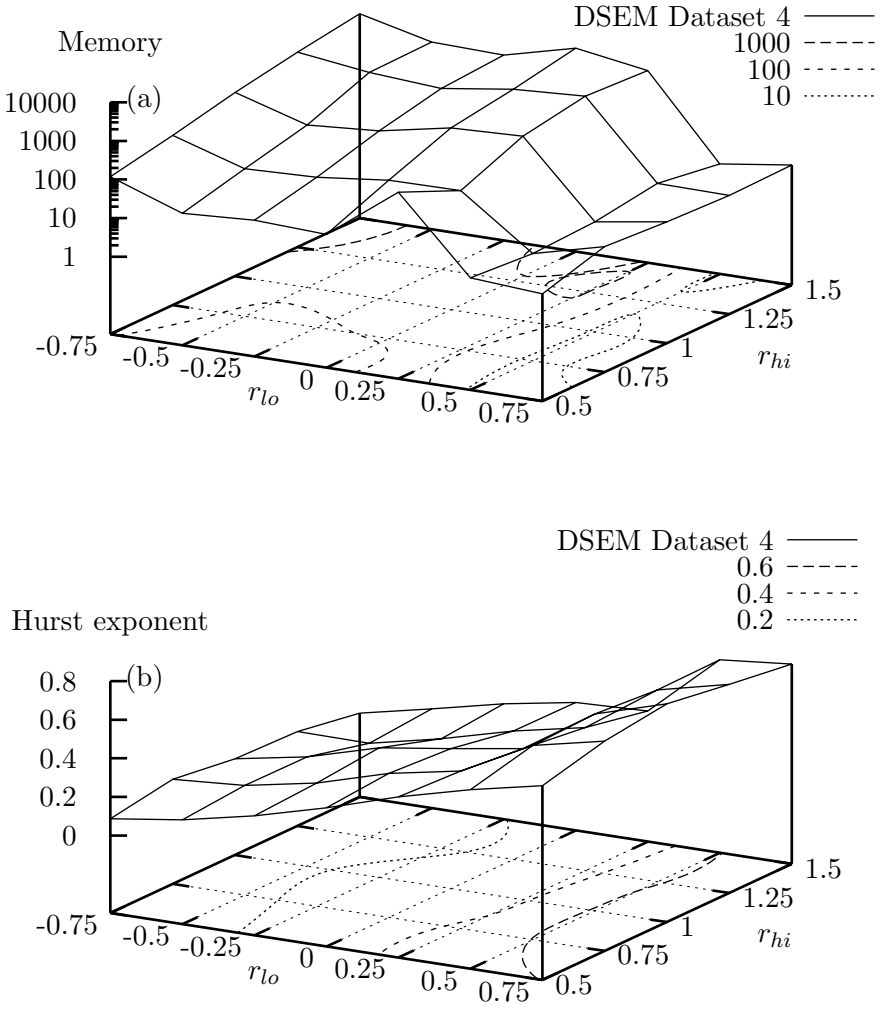

Figure 5.14: Both the crossover point, or memory, (a) and Hurst exponent for short timescales (b) indicate that memory effects are minimized when $r_{lo} \geq 0.25$ in DSEM with a two-point price response distribution. (The high values of the Hurst exponent for $r_{lo} > 0.5$ (b) do not cause problems because the memory is very short in this region (a).)

representative of a real market, the question is naturally raised, "Why do the agents choose this region of parameter space?" The simplest (but unjustified and hardly satisfactory) explanation is simply that any other choice would provide arbitrage opportunities which could be taken advantage of by watching for trends in the price series. So a rational agent would choose a nonzero price response in the expectation of these arbitrage opportunities, but ironically, in doing so, the memory (and opportunities) disappear! In other words, expectations of information in the price series erase that very information.

5.3 Volatility clustering

In the last section we observed that the price series in DSEM always crossed over to a domain with no memory effects for long enough times. This would seem to imply a lack of history-dependence in the time series. However, it has been empirically observed that volatility (to be defined) has a *very* long memory. This leads to the phenomenon of clustered volatility: high activity in the market is observed to cluster together, separated by spans of low activity.

The simplest definition of volatility is the absolute value of the price return over some interval. (This definition appears to be more prevalent than the square of the returns [36, 77, 78, 80].) Clustered volatility, by this definition, means there exists periods in which the price changes rapidly and dramatically, separated by other periods where few/small changes in the price occur. Hence, there exist long-range temporal correlations in the absolute value of price returns.

As before, the Hurst exponent is a promising quantity to measure these correlations. For the daily returns of the Dow Jones Industrial Average [3] shown in Fig. 5.10, the Hurst exponent of the volatility is measured to be $H = 0.852 \pm 0.009$ demonstrating very strong positive correlations—high volatility tends to be followed by further high volatility and low by low. Other studies measuring the Hurst exponent of the volatility from empirical data have concluded $H \approx 0.9$ [77], $0.63 \leq H \leq 0.95$ [78], and $H \approx 0.85$ [36]. Two other works I am aware of calculated the exponent of the autocorrelation function which decays as a power law with exponent $2H - 2$ [82], giving $H \approx 0.9$ [80] and $H \approx 0.8$ [4]. (The latter defined volatility as the squared return rather than its absolute value, but came to the same conclusions.) These studies were performed on a variety of systems so, clearly, clustered volatility is universal.

Again, we seek to know whether DSEM also exhibits this property. To test it, we continue with our analysis on a tickwise (number of trades executed) timescale and calculate the Hurst exponent for the absolute value of the returns. The

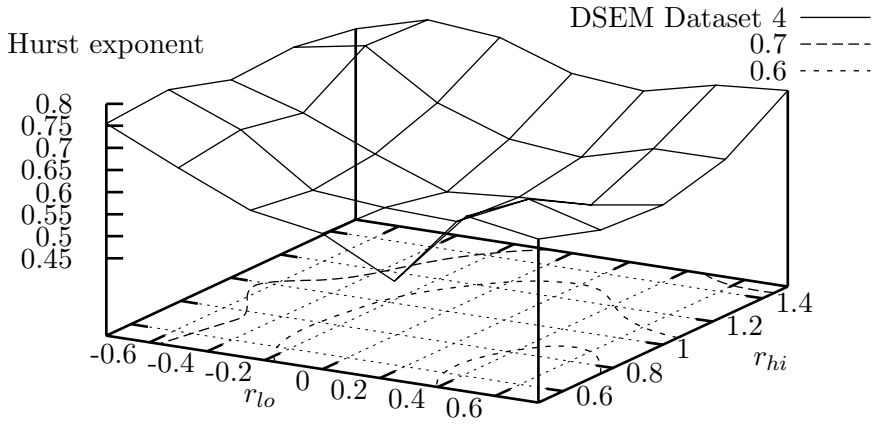

Figure 5.15: The Hurst exponent of the absolute returns, which measures the degree of clustered volatility, is strictly greater than one half for all parameter combinations in DSEM. It is particularly high when the upper limit of the two-point distribution r_{hi} is large or when the lower limit r_{lo} is small.

results are summed up in Fig. 5.15 which shows that the Hurst exponent measuring volatility clustering is always above one half over the whole parameter space, but significantly so when the upper limit r_{hi} of the two-point price response distribution is large or the lower limit r_{lo} is small.

Overall, though, the Hurst exponents are somewhat smaller (the greatest value was $H = 0.77$) than the empirical results ($H \approx 0.9$), suggesting that our search space should be expanded to larger values of r_{hi}.

5.3.1 Shuffling

As a check of the analysis the absolute return data for a particular run ($r_{lo} = 0.75$, $r_{hi} = 1.5$ with $H = 0.72 \pm 0.02$) were shuffled and the Hurst analysis of the shuffled data recalculated. Shuffling destroys temporal correlations so the expected value of the exponent is one half. In fact, the Dow Jones Industrial Average data yields $H = 0.502 \pm 0.010$ for the absolute returns when shuffled. For the sample DSEM data the resultant exponent is $H = 0.520 \pm 0.006$. Both are very close to one half, confirming that the high value for the unshuffled absolute-return data is due to temporal correlations (clustered volatility).

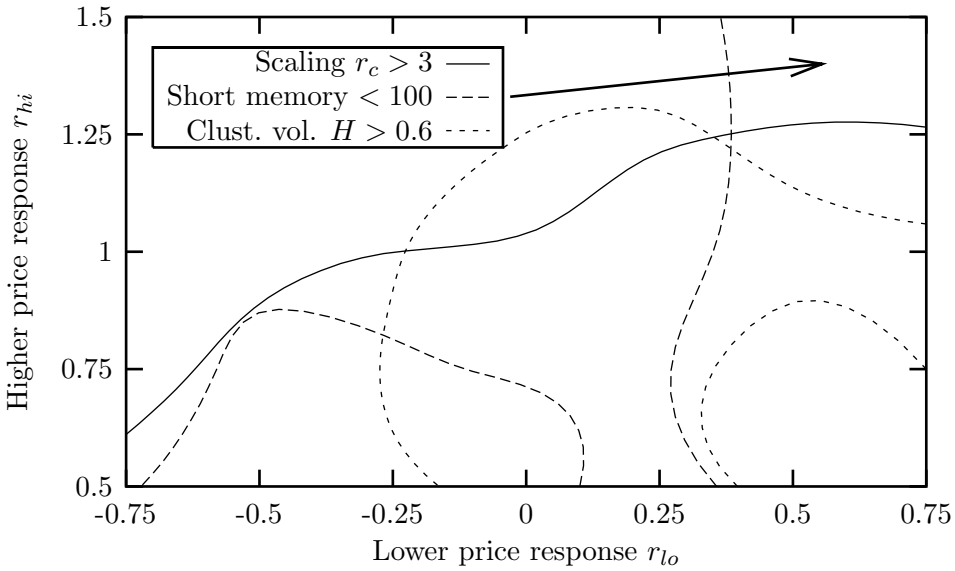

Figure 5.16: DSEM Dataset 4 ($N = 100$ agents) is able to capture three important properties observed empirically when $r_{lo} > 0.35$ and $r_{hi} > 1.25$. The curves are contours from previous plots: (1) characteristic return $r_c = 3$ from Fig. 5.8 (solid line); (2) memory in return series = 100 from Fig. 5.14(a) (dashed line); and (3) Hurst exponent for the absolute returns $H = 0.6$ from Fig. 5.15 (dotted line).

5.4 Scaling and Clustered volatility

In real markets all three properties of (1) scaling, (2) uncorrelated returns, and (3) clustered volatility are observed. As we have seen, DSEM can replicate each of these features when the agents have a two-point price response (r_{lo} and r_{hi}) and the values of these parameters are chosen appropriately. Of particular interest is whether there is a region of parameter space in which all three of these phenomena are observed simultaneously.

In some of the previous plots contour lines were drawn to indicate separation into regions which did and did not exhibit the phenomenon of interest: Fig. 5.8 plotted the contour line $r_c = 3$ to distinguish between parameter combinations which did ($r_c > 3$) and did not exhibit scaling in the return distributions. Fig. 5.14(a) separates parameter combinations that do not have a long memory (< 100) in the return series from those that do. Finally, Fig. 5.15 measures, with the Hurst exponent, the clustered volatility where $H > 0.6$ indicates the presence of clustered volatility while $H < 0.6$ indicates its absence.

These three contour curves are plotted together in Fig. 5.16 showing that all

three empirical properties are only observed in the upper right corner of the graph, when $r_{lo} > 0.35$ and $r_{hi} > 1.25$. So DSEM is most realistic for these parameter combinations.

One point to note regarding this region is that it spans the critical point at $r_p = r_1 = 1$ with the low end of the distribution below and the high end above. It is also interesting that the value $r_{lo} = 0$ (characterizing agents that do not use the return series as an indicator of performance) is not in the "realistic" region.

5.5 Wealth distribution

Thus far we have explored only the temporal dynamics of the stock price. But the distribution of wealth among agents may be of interest as well. It is well known that incomes in many populations are distributed log-normally [83] (with a particular exception we will come to later). So it is natural to ask how wealth is distributed in DSEM. (Again, we neglect the analysis of CSEM.)

5.5.1 Challenges

Determining the wealth distribution in DSEM is problematic because long runs of many agents are required. The durations must be long because the wealth distribution is initialized to a delta function (initially, all agents have the same amount of cash and stock) so a long transient is to be expected before any steady-state emerges.

But this constraint must be balanced against the need for many agents. Most of the simulations presented in this thesis consisted of $N = 100$ interacting agents—a rather small number for any statistical description. But computational limitations prevent serious investigation of larger systems because the number of operations grows as N^2. So we must be content with the data we have collected so far.

5.5.2 Log-normal distribution

A rigorous analysis of the distribution of wealth will not be attempted. Instead, it is merely reported that a log-normal distribution was suitable in most cases for Datasets 1–3. However, Fig. 5.17—showing a typical distribution—highlights the difficulties in establishing the proper distribution: the small agent numbers combined with the narrow range of wealths observed allows one only to say that the distribution is unimodal, but nothing more.

Even so, simply knowing that the distribution is unimodal is satisfactory. Any other result would be surprising because the agents in Datasets 1–3 are all

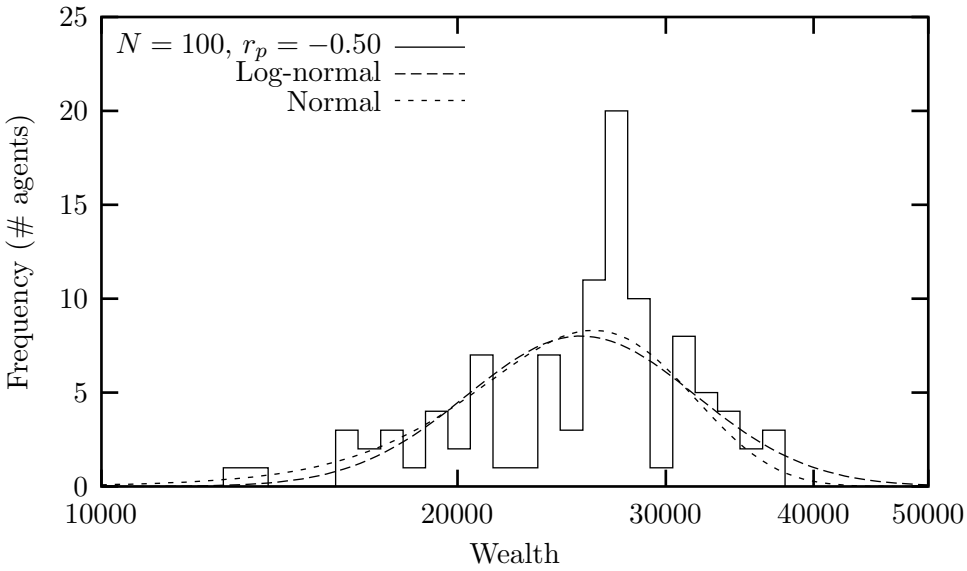

Figure 5.17: Sample distribution of agents' wealth from DSEM Dataset 3 ($N = 100$, $r_p = -0.50$). There is insufficient data to distinguish between a normal and a log-normal distribution.

of a similar character, varying (continuously) only in their news responses r_n and frictions f, as shown in Table 4.4.

5.5.3 Two-point price response

Worth further investigation are the data from Dataset 4 (Table 5.2) where the agents varied *discontinuously* in their price responses. In this case the market consists of two distinct populations with fundamentally different behaviours, so one would not expect the wealth to be distributed unimodally. A reasonable alternative is that each population has a log-normal wealth distribution producing a bimodal distribution over the whole market.

A representative distribution is shown in Fig. 5.18, confirming the bimodal hypothesis. Interestingly, the agents with $r_p = r_{lo} = 0$ outperform (have more wealth) than their $r_p = r_{hi} = 1$ counterparts.

To determine the generality of this result the average wealth for each sub-population was computed for all the simulations in Dataset 4. The hypothesis that the sub-population with the smallest absolute price response $|r_p|$ would outperform the more reactive agents was tested by comparing the difference between the absolute values of the responses $|r_{hi}|-|r_{lo}|$ and the ratio of wealth held by each sub-population ($w(r_{hi})$ vs. $w(r_{lo})$). If valid then we should find the ratio of the wealths obey

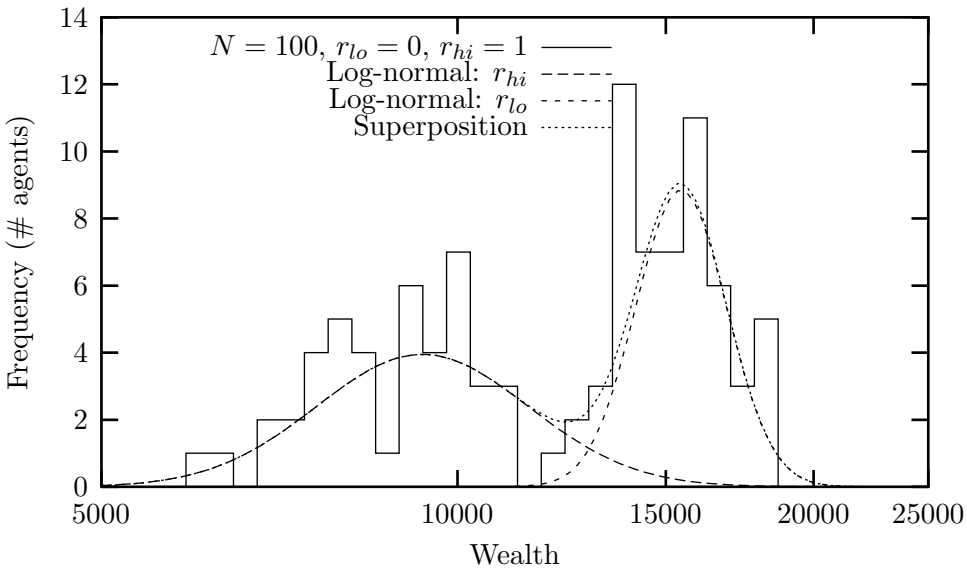

Figure 5.18: Sample distribution of agents' wealth from DSEM Dataset 4 ($N = 100$, $r_{lo} = 0$, $r_{hi} = 1$). The log-normal curves are calculated from each sub-population, revealing a strongly bimodal nature.

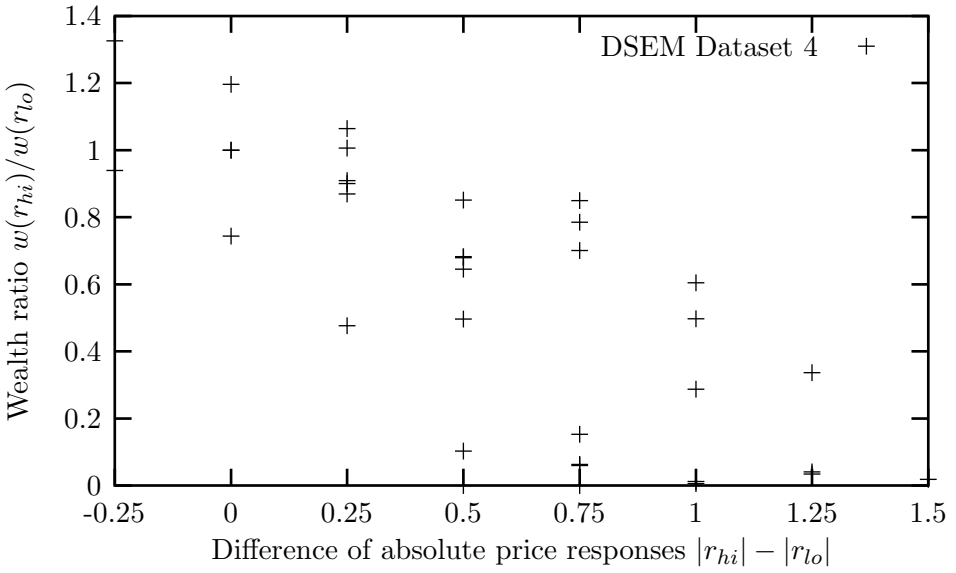

Figure 5.19: In DSEM with a two-point price response the wealth of each of the sub-populations $w(r_p)$ depends strongly on the magnitude of the price response $|r_p|$. The population with the smallest absolute price response (r_{hi} to the left of zero and r_{lo} to the right) consistently has more wealth as indicated by the ratio of wealth between the two sub-populations.

$w(r_{hi})/w(r_{lo}) > 1$ when $|r_{hi}| < |r_{lo}|$ and vice versa. Fig. 5.19 demonstrates this hypothesis holds very well—with a linear correlation for the plotted data of –76%—suggesting that the "best" strategy is to ignore the price fluctuations, $r_p = 0$.

This raises an interesting question: if a zero price response is best, why would agents choose non-zero values? We have seen in previous sections that realistic market phenomena such as scaling and clustered volatility only emerge in when the price responses are set far from zero. If DSEM is meant to represent real investor behaviour, why do investors base their decisions on price fluctuations when the model indicates this is detrimental?

This issue is currently being investigated by allowing the agents to "learn" from their past mistakes as will be discussed in Section 7.4. The purpose of this research is to determine if nonzero values of the price response parameter emerge spontaneously in the dynamics via the learning process.

Pareto's law of income distribution

In 1897 V. Pareto noticed that incomes tend to be distributed log-normally over the majority of the sample data excepting the tail of the distribution (the highest one percent of the incomes) which decay as a power law, an observation which holds in many countries to this day [83]. CSEM and DSEM were deliberately constructed to keep a record of agents' wealths so that this claim could be tested. However, testing for this property requires even larger system sizes than we have explored so far, since the highest percentile of a population of even $N = 1000$ consists of only ten agents—inadequate for statistical analysis.

Larger systems are currently under investigation but insufficient data were available as of completion of this dissertation.

5.6 Summary

In this chapter a number of unusual qualities of empirical markets were explored in the context of the Centralized and Decentralized Stock Exchange Models (CSEM and DSEM, respectively). CSEM was unable to reproduce even the first of these: scaling in the tail of the price return distribution. DSEM was also unable to produce this effect until the restriction that all agents maintain the same price response parameter r_p was relaxed and instead two values were allowed—r_{lo} and r_{hi}—thereby splitting the population into two distinct "types." Then, for particular values of the parameters, scaling was observed over a range of returns in excess of three standard deviations.

Two other empirical phenomena were also explored: the lack of correlations in the return series but the presence of long-range correlations in the volatility (defined as the absolute value of the return). Having failed the first test CSEM was not tested but DSEM was able to capture both these properties, again in a suitable region of parameter space.

All three of these properties were observed simultaneously in DSEM when $0.5 \leq r_{lo} < 1$ and $r_{hi} \geq 1.25$, spanning the critical point at $r_p = r_1$ discovered in Section 4.2.4. The significance of this result is unclear but some thoughts on the matter are discussed in Chapter 7. But first the results of some interesting experiments with some real stocks are discussed.