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 4

Analysis and Results: Phase space

In the previous two chapters the Centralized and Decentralized Stock Exchange Models (CSEM and DSEM, respectively) were presented and in each case all but two parameters were fixed. In this chapter the remaining parameter space will be investigated and it will be demonstrated that both models exhibit phase transitions for interesting values of these parameters. We begin with CSEM.

4.1 CSEM phase space

4.1.1 Review

The Centralized Stock Exchange Model (CSEM), presented in Chapter 2, consists of a number N of agents which trade once daily with a centralized market maker. The market maker chooses a trading price such that all orders are satisfied and the market clears (supply exactly balances demand). The agents choose their orders based on a forecast of the daily return-on-investment which has a stochastic component modeled as a Gaussian deviate with standard deviation σ_ϵ (defined as the forecast error). In Chapter 2 the model parameter space was reduced leaving only N and σ_ϵ as free parameters. In this section the remaining two-dimensional parameter space will be explored.

4.1.2 Data collection

To explore the phase space thoroughly simulations were performed on systems of sizes N=50, 100, 200, 500, and 1000 with forecast errors in the range $\sigma_\epsilon \in [0.01, 0.50]$ for each N, with increments of 0.01 up to $\sigma_\epsilon = 0.25$ and increments of 0.02 thereafter,

Parameters	CSEM Dataset 1				
Particular values					
Number of agents N	50	100	200	500	1000
Investment limit δ	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}
Run length (time steps)	10,000	10,000	20,000	20,000	30,000
Number of runs	38	38	38	38	38
Common values					
Forecast error σ_ϵ	0.01 to 0.25 by 0.01 0.26 to 0.50 by 0.02				
Total cash C	\$1,000,000				
Total shares S	1,000,000				
Memory M	105 ± 95 (uniformly distributed)				
Risk aversion a	2 ± 1 (uniformly distributed)				
Degree of fit d	0 (moving average)				
seed	random				

Table 4.1: Parameter values for CSEM Dataset 1. Some of the parameters were established in Chapter 2 and are common to all the runs. Dataset 1 explores two dimensions of phase space: N and σ_ϵ.

for a grand total of 190 experiments performed. The complete list of parameter values used can be found in Table 4.1.

The choices of parameter values used (other than N and σ_ϵ) are justified in Chapter 2. To introduce heterogeneity amongst the agents some of the parameters, namely the memory M and risk aversion a, were chosen randomly for each agent from the ranges indicated in the table (with the deviates uniformly distributed within the ranges).

Each run consisted of at least 10,000 time steps (days) and larger systems had longer runs to compensate for slower convergence to the steady state. (With these run lengths the initial transient never accounted for more than one third of the total run).

4.1.3 Phases

In most of the runs an initial transient period was observed before the price converged to a steady state value around which it fluctuated. The only discrepancy was for small forecast errors where the price climbed quickly until it reached a maximum value which it often returned to. Representative plots of these behaviours are shown in Fig. 4.1(a) and (b), respectively.

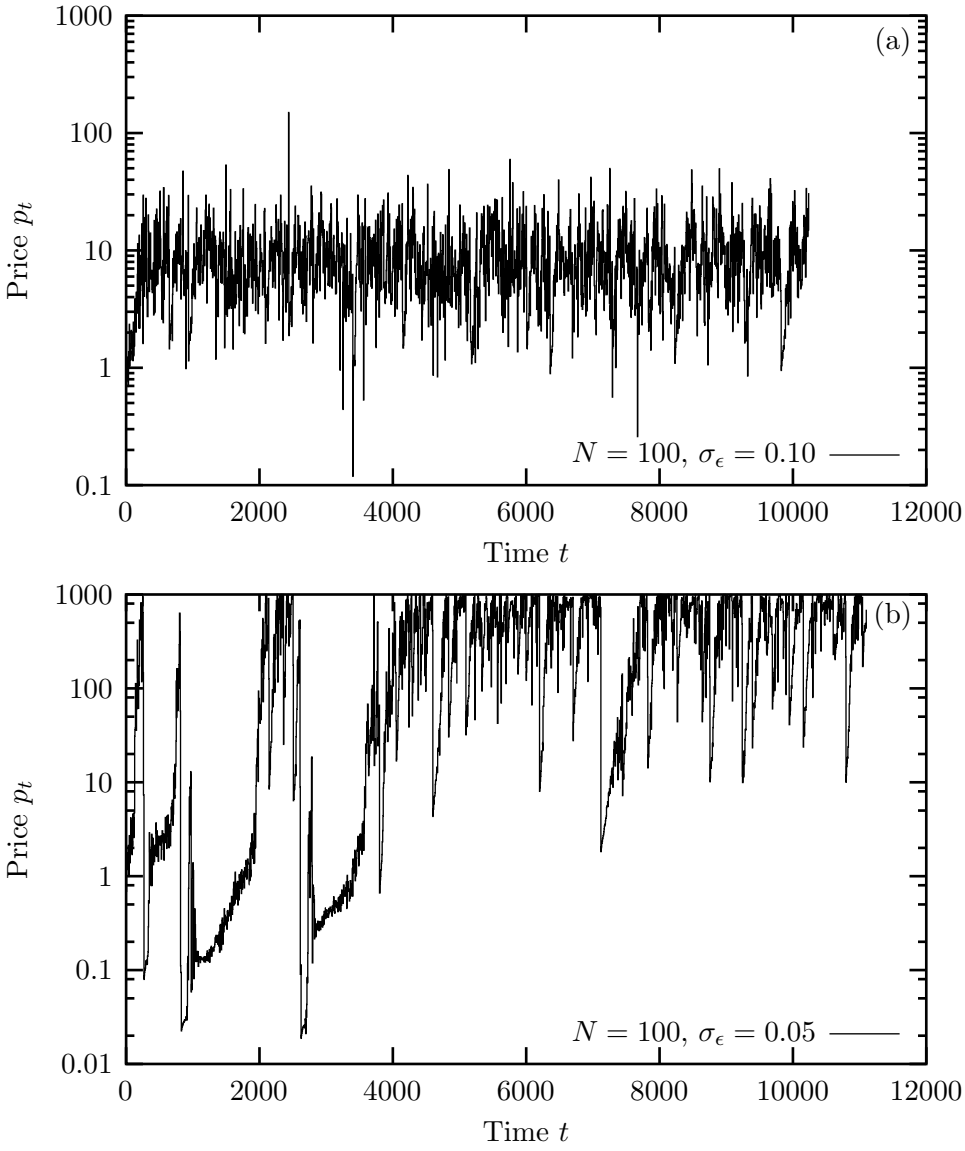

Figure 4.1: The price series plots for CSEM with $N = 100$ agents and $\sigma_\epsilon = 0.10$ (a) and $\sigma_\epsilon = 0.05$ (b) indicate a change of character of the dynamics.

Parameters	CSEM Dataset 2		
Number of agents N	100	100	100
Investment limit δ	10^{-2}	10^{-4}	10^{-5}
Number of runs	38	38	38
Run length (time steps)	10,000	10,000	10,000

Table 4.2: Parameter values for CSEM Dataset 2. These runs are a variation of Dataset 1 (all unspecified parameters are duplicated from Table 4.1, $N = 100$) exploring a range of investment limits δ.

The transition between these two behaviours was observed for all system sizes near $\sigma_\epsilon \approx 0.08$ and is most dramatic when looking at the maximum price observed in a run.

Maximum price

As is demonstrated in Fig. 4.1(a) the price in each run converged to some steady-state value after some time and then appeared to randomly fluctuate around that value, never exceeding some maximum. As mentioned above, the only exception was when the price reached a limit which interfered with its natural fluctuations.

The limit price is a consequence of the investment limit parameter δ introduced in Section 2.4.3 where it was noted that the price may not exceed the limit $p_{max} = (1 - \delta)/\delta$ (Eq. 2.44).

Fig. 4.2 clearly captures the distinct character of the dynamics on both sides of $\sigma_\epsilon \approx 0.08$. For larger σ_ϵ the price fluctuates freely while for smaller values the limit has a strong influence on the dynamics.

In the limit $\delta \to 0$ (which is disallowed because it can occasionally generate singularities in the price series) it appears that the maximum price would diverge at $\sigma_\epsilon = 0$ producing a phase transition. Although this cannot be tested by directly setting $\delta = 0$ the limit can be explored by studying smaller values of δ.

4.1.4 Investment limit

A subset of Dataset 1 with N=100 agents was simulated again, but this time with $\delta = 10^{-2}$, 10^{-4} and 10^{-5} as shown in Table 4.2. It was suspected that reducing δ would increase the limit price and thereby allow the price to fluctuate freely for smaller values of σ_ϵ, reducing the domain of the second phase.

However, as Fig. 4.3 demonstrates the threshold values of σ_ϵ (see Table 4.3)

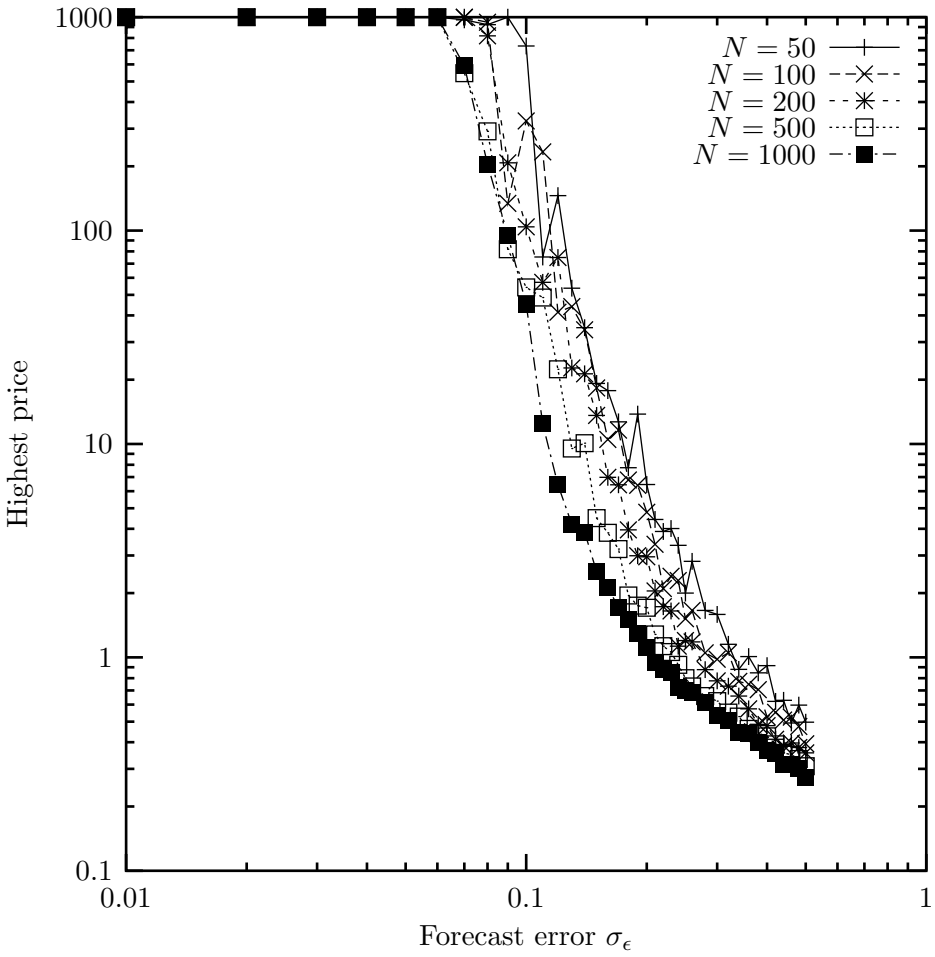

Figure 4.2: The highest price in any given simulation increases as the forecast error decreases until it reaches its theoretical limit, creating two separate phases for the dynamics.

δ	Threshold σ_ϵ
10^{-2}	0.08
10^{-3}	0.06
10^{-4}	0.06
10^{-5}	0.08

Table 4.3: The threshold values of σ_ϵ separating the two phases of CSEM shown in Fig. 4.3 do not appear to depend on the investment limit δ.

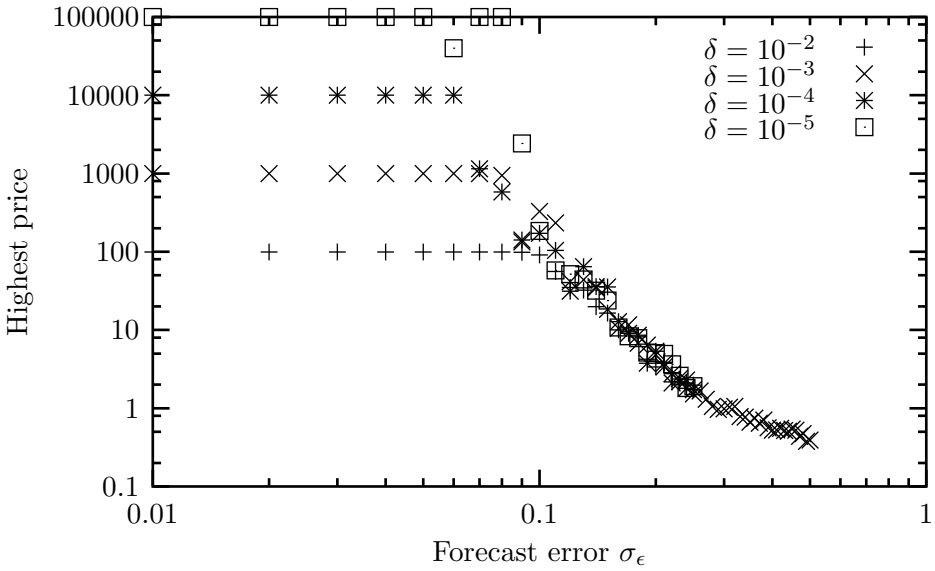

Figure 4.3: The maximum price in CSEM has a limit which depends on the investment limit δ. However, the threshold value of σ_ϵ for which the limit is first reached does not appear to depend on δ.

for which the price first reaches its limit remains constant, even though the price limit increases. This suggests that there exists a critical forecast error $\sigma_c > 0$ at which the maximum price diverges. Even though the critical point is only strictly defined in the limit of $\delta \to 0$ the term will also be used here to refer to systems with nonzero values of δ.

4.1.5 Critical regime

Critical points are heralded by power law relationships of the form $f(x) \propto (x - x_c)^z$ where x_c is the critical point and z is known as the *critical exponent*. (For a thorough explanation of critical phenomena see Ref. [61].) Many different quantities can play the role of the critical variable f. In a thermodynamic system it could be an order parameter such as the magnetization of a ferromagnet. Alternatively, f can be a response function such as the susceptibility or the specific heat or it could be a correlation time or time for thermalization. The control parameter x could be the temperature or an external field. In the case of CSEM we will continue to use the maximum price as the order parameter and σ_ϵ as the control parameter.

To specify the transition in more detail it would be helpful to estimate the critical point σ_c and the exponent from the data. We begin by reconsidering the

data from Fig. 4.3 and fitting it to a power law

$$p_{max} = C(\sigma_\epsilon - \sigma_c)^{-b} \tag{4.1}$$

to estimate σ_c and b (C is unimportant).

The fitting algorithm used is a Levenberg-Marquardt nonlinear routine [20, Section 15.5] and the fit is performed over the range $\sigma_\epsilon \geq 0.14$. (Choosing a range which is too near the actual critical point tends to reduce the quality of the fit because critical points tend to be "blurred" on finite systems.) The fitting algorithm attempts to minimize the sum-of-squares error between the curve and the data but it can get stuck in suboptimal solutions which depend on the initial parameter choices when performing the fit. For these fits the parameters were initially set to $C = \exp(-3)$, $\sigma_c = 0.08$, and $b = 2$ because these values were observed to fit the data reasonably well. (Though setting $\sigma_c = 0.01$ initially, the fit still converged to the same solution.)

The resultant fits for each value of δ in Dataset 2 (all with $N = 100$) and for $N = 100$ in Dataset 1 give the critical points and exponents shown in Fig. 4.4. The weighted average of the exponents is $b = 1.57 \pm 0.10$ and the mean critical forecast error is $\sigma_c = 0.120 \pm 0.005$. The fact that these values are similar for all values of δ tested strengthens the conclusion that σ_c is a critical point.

Notice that the calculated value of σ_c is significantly higher than the 0.08 originally hypothesized. This is a common feature of experiments involving critical phenomena and is due to the finite size of the system under investigation. A true critical or *second-order* phase transition is characterized by a discontinuity in the derivative of the order parameter. In finite systems the discontinuity is smeared out and becomes more refined with larger systems. In this case the smearing resulted in an inaccurate first guess of the critical point. After exploring some alternative choices for the order parameter we will consider finite size effects in more detail.

4.1.6 Alternative thermodynamic variables

In the last section the maximum price over any run was chosen as the thermodynamic property whereby the phase transition was detected. In this section we demonstrate that a number of alternative variables would be equally suitable.

In particular, we consider two alternatives: the (logarithmic) average of the price series

$$\bar{p} \equiv \exp \langle \log p \rangle \tag{4.2}$$

and the (logarithmic) variance around the average

$$\sigma_p^2 \equiv \mathrm{Var} \left[\log p \right]. \tag{4.3}$$

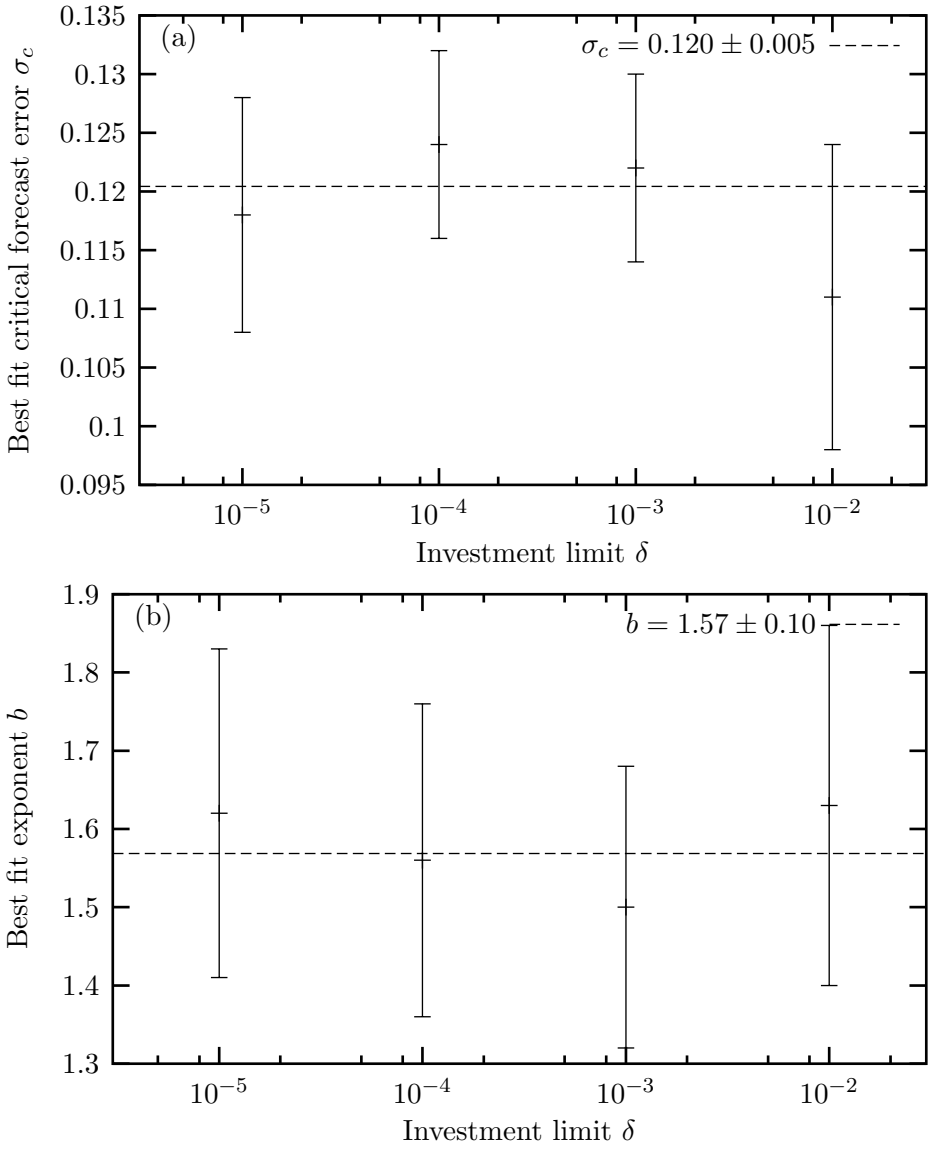

Figure 4.4: The best fits of power laws to CSEM Dataset 2 (and $N = 100$ from Dataset 1) yield the critical points (a) and scaling exponents (b) shown. The lines represent the weighted averages of the best fit values.

The variance σ_p^2 measures the scale of the fluctuations and is analogous to magnetic susceptibility in non-equilibrium systems.

Notice a phase transition in the average price $\bar{p}$ is equivalent to one in the (time-averaged) wealth per agent $\bar{w}$ because they are related by

$$N\bar{w} = C + \bar{p}S \tag{4.4}$$

where C and S are the total amount of cash and shares, respectively.

Fig. 4.5 demonstrates that both these properties exhibit scaling, diverging at the critical point $\sigma_c = 0.12$ (from Fig. 4.4(a)). The critical exponent for the average price power law is 1.11 ± 0.04 and the exponent for the fluctuations is 0.94 ± 0.02. Clearly these properties would be equally suitable to determine the phase transition, but the maximum price has the advantage that the transition becomes very clear because it is a constant to the left of the critical point, being bounded by δ.

As mentioned before, the deviation from scaling observed in the variance plot is due to finite size effects which we explore next.

4.1.7 Finite size effects

Now we return to the maximum price data in Fig. 4.2 and fit it to a power law using the technique described before. The resultant estimates of the critical point are shown in Fig. 4.6(a) which demonstrates that as the system size N increases the critical point decreases systematically (neglecting the smallest system which is plagued by noise). To derive the relationship between the system size and the associated critical point we need to understand the role of correlations.

Correlations

Near a critical point the dynamics are dominated by correlations between elements of the system (agents, in our case). The degree to which the elements are correlated is measured by the correlation length ξ which, in CSEM, counts the typical number of agents affected by any single agent's decision. Far away from the critical point we don't expect one agent's decisions to affect (many) other agents so the correlation length is short. But near the critical point the correlation length diverges as [61]

$$\xi(\sigma_\epsilon - \sigma_c) \propto (\sigma_\epsilon - \sigma_c)^{-\nu}. \tag{4.5}$$

When dealing with a finite system the correlation length is attenuated by the size of the system N. It should reach a maximum at the critical point for that particular system size, denoted by $\sigma_c(N)$, and we expect the maximum to grow linearly with N,

$$\xi(\sigma_c(N) - \sigma_c) \propto N. \tag{4.6}$$

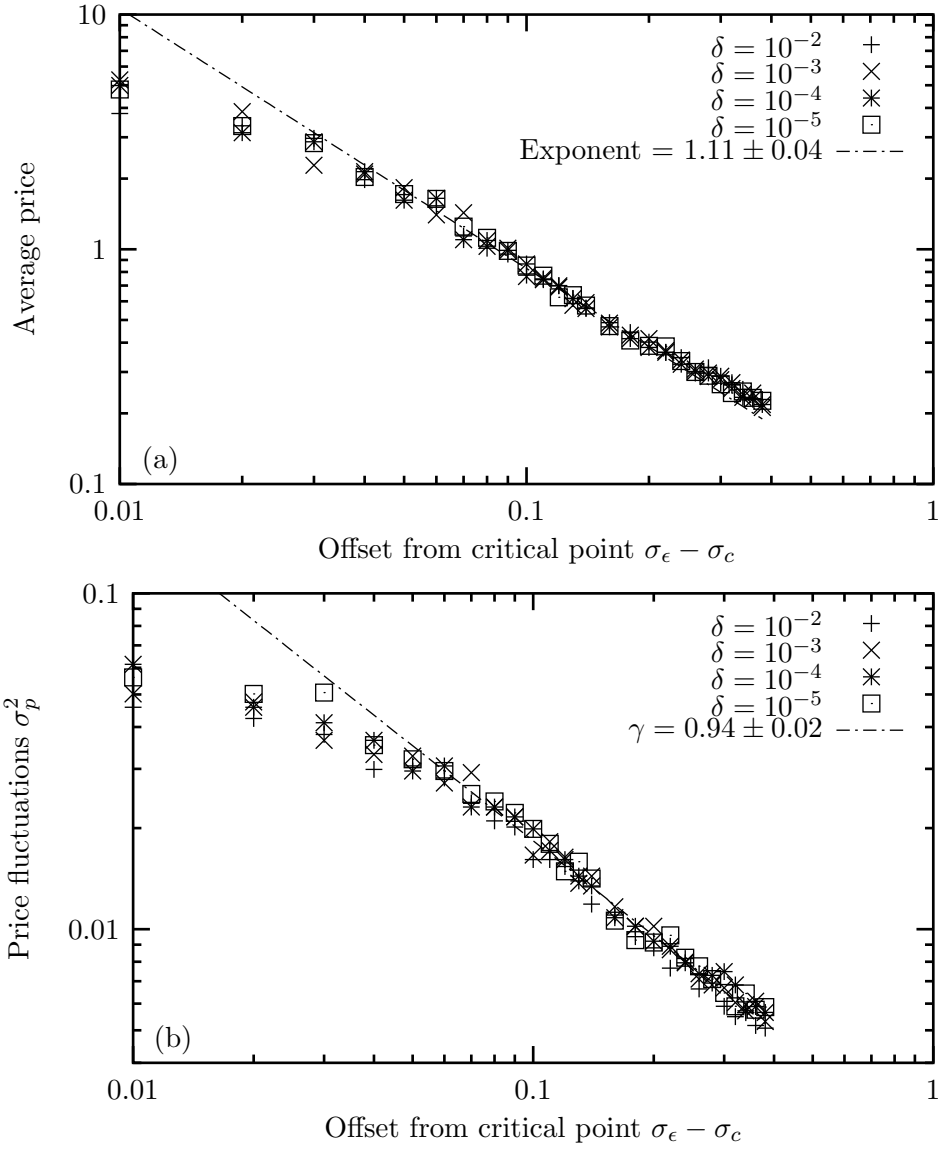

Figure 4.5: The average price (a) and variance of fluctuations (b) also exhibit scaling near the critical point $\sigma_c = 0.12$ for the data from CSEM Dataset 2. The deviation from scaling observed near the critical point in (b) is due to the finite size of the system ($N = 100$) as will be seen in Fig. 4.7.

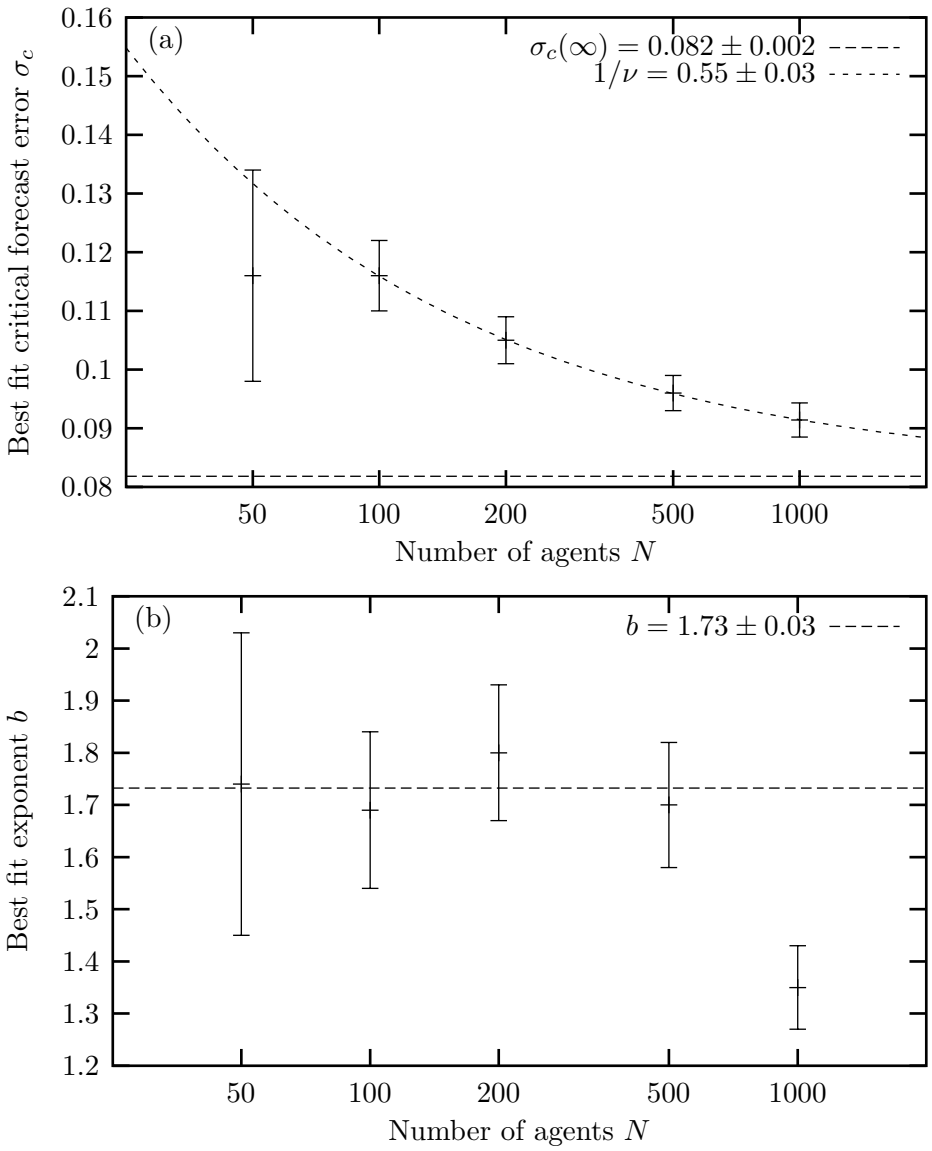

Figure 4.6: The best fits of power laws to CSEM Dataset 1 yield the critical points (a) and scaling exponents (b) shown. A finite-size scaling analysis (neglecting $N = 50$) reveals information on how the critical point changes with increasing investor numbers (a). For reasons discussed in the text, the exponent for $N = 1000$ is dropped from the estimate of the scaling exponent (b).

As a consequence of these two equations we expect the finite-size critical point to converge to the thermodynamic critical point as

$$\sigma_c(N) - \sigma_c \propto N^{-1/\nu}. \tag{4.7}$$

Applying this relationship to the data in Fig. 4.6(a) gives a finite-size scaling exponent $1/\nu = 0.55 \pm 0.03$ which means that the correlation length grows as $\xi \sim (\sigma_\epsilon - \sigma_c)^{-\nu}$ with $\nu = 1.82 \pm 0.10$. It also allows a more precise estimate of the (limit) critical point, giving $\sigma_c = 0.082 \pm 0.002$.

Scaling exponent

The best estimates of the critical exponent b from Eq. 4.1 are shown in Fig. 4.6(b), giving an average value of $b = 1.73 \pm 0.03$. Notice the largest system size gives a markedly different result and is not used to compute the average b. This is a consequence of the range of forecast errors over which scaling applies:

Notice that the tails of the highest prices ($p_{max} < 1$) in Fig. 4.2 appear slightly "flatter" than the rest of the data. In fact the tails fit quite well to the scaling relation $p_{max} \propto \sigma_\epsilon^{-1}$ which is probably directly related to the inverse relationship derived in Section 2.2.9. This inverse power law can obscure the critical scaling so the range over which the scaling was tested for was constricted to $p_{max} \geq 1$.

Unfortunately, since the run for $N = 1000$ had the lowest observed prices at any given value of σ_ϵ this restriction severely limited the available data and compromised the fit. It appears that the tail is artificially drawing the estimate of b down for this system size, so it was not included in the computation of $b \approx 1.73$.

Fluctuations

In this section we briefly revisit the scaling seen in the fluctuation data (Fig. 4.5(b)) to demonstrate the round-off seen near the transition is a finite-size effect.

Since the fluctuations in the log-price series are due to stochasticity in each of the N agent's trading decisions it is reasonable to expect the variance of the log-price to scale with system size as $\sigma_p^2(N) \propto 1/N$ so that $N\sigma_p^2$ should be independent of system size. For the most part Fig. 4.7 confirms this hypothesis with some deviations near the critical point. Notice these deviations diminish with larger system sizes so these deviations are just finite-size effect—a "blurring" of the phase transition for small system sizes.

The best estimate of the scaling exponent, taken from the largest system is $\gamma = 1.29 \pm 0.02$.

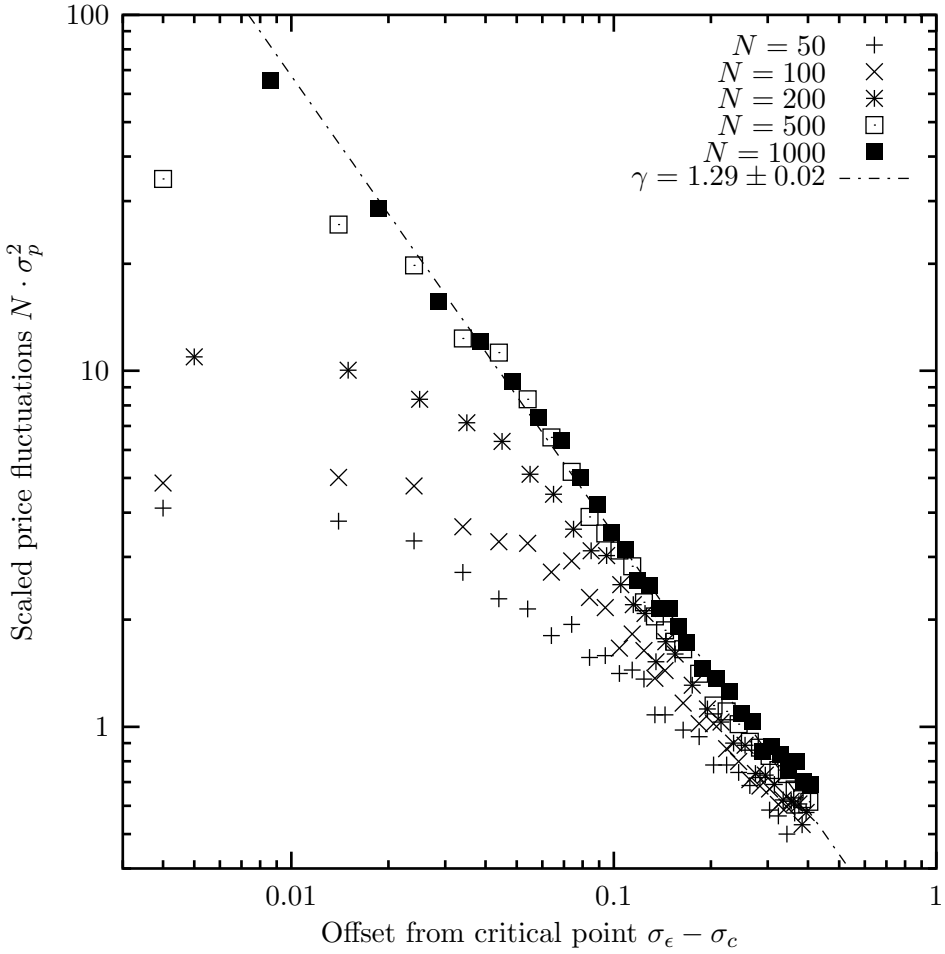

Figure 4.7: The variance of the log-price largely collapses to a single curve when multiplied by the system size N for CSEM Dataset 1. This curve diverges as the critical point is approached with an exponent $\gamma = 1.29 \pm 0.02$ calculated from the largest system $N = 1000$. (The critical points were taken from Fig. 4.6(a).)

Universality class

The main advantage of characterizing the critical point precisely is that the critical exponents b, γ, and ν may tell us to which *universality class* the critical point belongs. At a critical point many of the particular details of a system become irrelevant and the scaling properties depend only on a few basic quantities, such as the dimensionality and symmetry of the system [61]. As such, many disparate systems are observed to behave in the same way at a critical point and may be classified by their common exponents.

Discovering which universality class a system belongs to leads to further understanding of the important features of the system. For instance, the author has recently been involved in research into the "game of Life" (GL), a toy model of interactions between spatially-distributed individuals. GL lies close to a critical point in the same universality class as directed percolation, a model of the spreading of a cluster through its nearest neighbours [14]. Making this connection teaches us that the important factor determining the dynamics at the critical point near GL is simply the probability of a disturbance spreading to its neighbours, not the particular details of GL.

I do not know which universality class CSEM belongs to, but by computing the exponents it is my hope that a reader will recognize them and classify the model. The exponent b probably is unrelated to physical systems but the variance of the fluctuations which gave $\gamma \approx 1.29$ is analogous to magnetic susceptibility. The exponent for the correlation length $\nu \approx 1.8$ may also be relevant even though the model is *mean field* (each agent interacts with all other agents through the market maker).

4.1.8 Transient

In this section we explore the critical phase transition discovered above from a different perspective; namely, that of the transient period. A close inspection of Fig. 4.1 will reveal that the price series had an initial transitory phase before it settled down to its steady-state dynamics. In this phase the memory of the initial conditions is slowly erased.

The transient can be systematically quantified by recognizing that the price series eventually converged to some steady-state value. Then the transient is formally defined as the period until the price first crosses its (logarithmic) average for the entire series.

The above definition is satisfactory provided that the steady state price is far from the initial price (on the order of \$1, see Fig. 2.3) but becomes irrelevant

for large σ_ϵ when the steady-state price is on the same scale as the initial price (see Fig. 4.2).

A plot of the transient duration of each run in Dataset 1, shown in Fig. 4.8, exhibits some interesting properties. For large forecast errors $\sigma_\epsilon \gg \sigma_c$ the transient measure is quite unstable—large for some runs and small for others—with some interesting system size dependence but these properties will not be analyzed further because the transient is a poor measure in this region.

For small σ_ϵ another interesting pattern is observed: the transient appears to grow near the critical point declining away from it, on both sides. This behaviour can be explained if the system does exhibit a second-order phase transition at σ_c. Criticality arises from correlations between agents which extend further and further as the system approaches the critical point. At the critical point the correlations span the entire system such that a perturbation in any element can have a cascade effect which may impact on any or all other elements. However, the correlations are an emergent phenomena in critical systems and require time to set up—the initial transient period. Away from the critical point the correlations do not span the entire system so they require less time to set up and, correspondingly, the transient is shorter.

Notice the maximum transient grows with the system size reflecting the longer time required for the correlations to span the system. Critical theory predicts the maximum transient should scale with the system size and Fig. 4.9 confirms it. The associated scaling exponent is estimated to be 1.5 ± 0.1. Since this exponent is greater than one the duration of the transient grows faster than N as the system size is increased—a typical property of criticality, known as *critical slowing down*.

4.1.9 Summary

In this section the phase space of the Centralized Stock Exchange Model (CSEM) was explored in detail. The main discovery was of a critical value of the forecast error $\sigma_c = 0.082 \pm 0.002$ above which the dynamics are relatively stable and below which the price fluctuates over many orders of magnitude, up to the maximum imposed by the investment limit δ. The transition exists for all values of δ explored ($10^{-5} \leq \delta \leq 10^{-2}$) and appears to be universal. Naturally, the transition becomes more pronounced for larger systems.

The economic interpretation of this transition is unclear. It is reasonable to expect the price to rise as uncertainty decreases, reflecting increasing confidence in the stock, but it is not obvious why the price would diverge for a non-zero uncertainty.

No other interesting phenomena were observed as the parameters were ad-

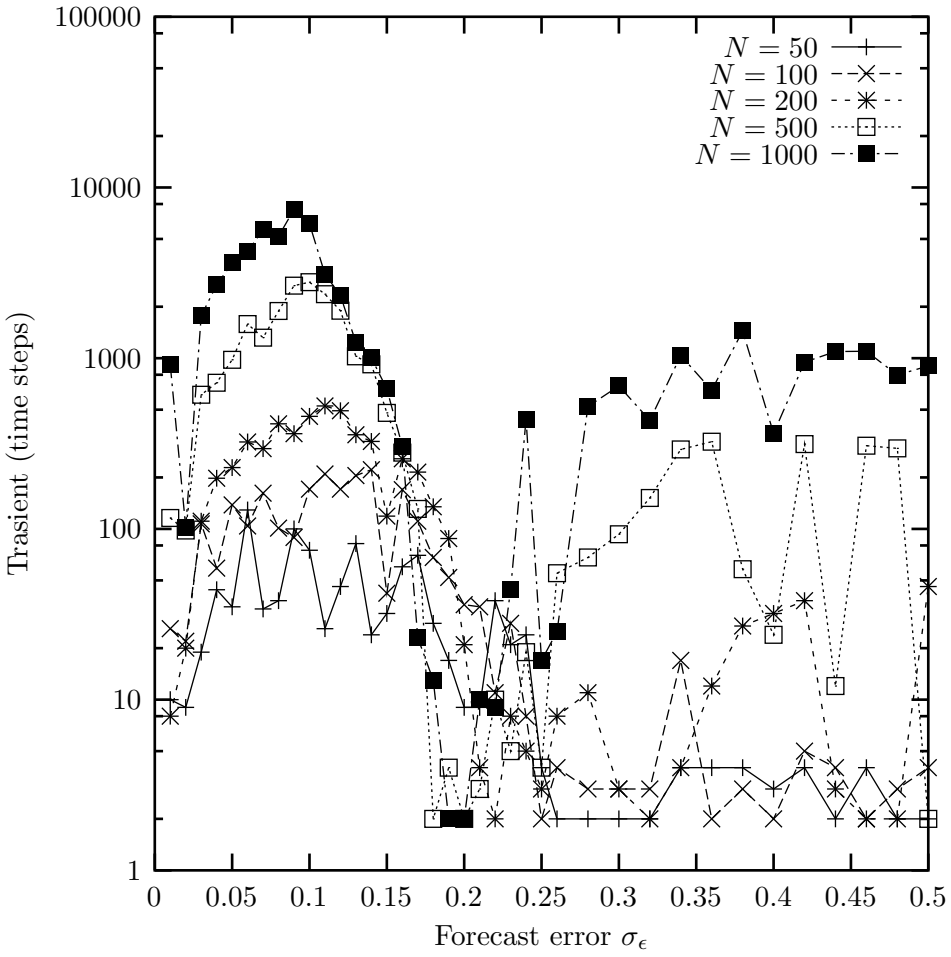

Figure 4.8: Duration of the transient period in CSEM (Dataset 1) before the price series settles down to some steady-state value. The transient grows near the critical point $\sigma_c \approx 0.08$.

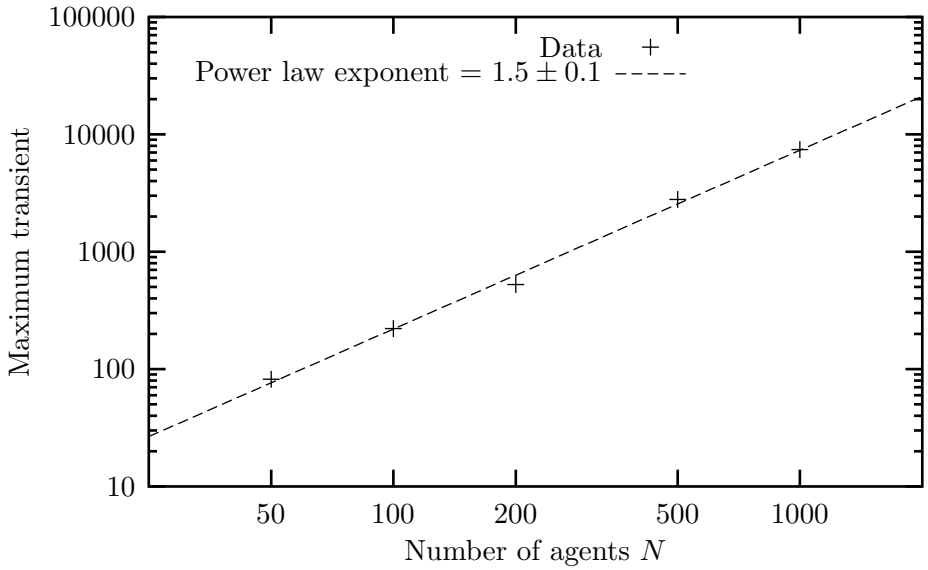

Figure 4.9: The maximum transient in CSEM appears to scale with the system size with an exponent 1.5 ± 0.1.

justed. In the next section, a similar analysis is performed on DSEM.

4.2 DSEM phase space

4.2.1 Review

The Decentralized Stock Exchange Model (DSEM), presented in Chapter 3, was constructed as an alternative to CSEM, discarding the notion of a centralized control which sets the stock price. In DSEM the price is an emergent property of agents placing and accepting orders with each other directly. The agents use a fixed investment strategy and place orders when their portfolios become unbalanced. The fraction of one's wealth each of the N agents keeps invested in the stock is affected by (exogenous) news and price movements. The degree of influence each of these factors has is parameterized by the news and price responsiveness, r_n and r_p, respectively. In Chapter 3 arguments were presented which reduced the parameter space, leaving only N and r_p as free parameters. In this section the role each of these parameters plays will be explored.

Parameters	DSEM Dataset 1				
Particular values					
Number of agents N	50	100	200	500	1000
Number of runs	39	39	39	39	39
Common values					
Price response r_p	–0.75 to 0.25 by 0.05 0.50 to 0.95 by 0.05 –0.34 to –0.31 by 0.01 0.91 to 0.94 by 0.01				
Total cash C	\$1,000,000				
Total shares S	1,000,000				
News interval τ_n	1				
News response r_n	0.01 ± 0.01 (uniformly dist.)				
Friction f	0.02 ± 0.01 (uniformly dist.)				
seed	random				
Run length ("days")	1,000				

Table 4.4: Parameter values for DSEM Dataset 1. Some of the parameters were established in Chapter 3 and are common to all the runs. Dataset 1 explores two dimensions of phase space: N and r_p.

4.2.2 Data collection

The phase space was explored by varying the price response parameter r_p and number of agents N. The choices of system sizes were N=50, 100, 200, 500, and 1000 agents while the price response was initially explored at a coarse resolution with increments of 0.25 between -0.75 and $+0.75$ then again with a finer resolution of 0.05 in the ranges -0.75 to $+0.25$ and 0.50 to 0.95. Finally, the regions $r_p \in [-0.34, -0.31]$ and $r_p \in [0.91, 0.94]$ were explored at a higher resolution of 0.01 because they exhibited interesting properties (to be discussed).

Although r_p is an agent-specific parameter it was set to a single value for all the agents reducing diversity somewhat. However, sufficient heterogeneity was maintained through the news response and friction parameters which were each spread over a range of values and each agent was randomly assigned a (uniformly distributed) deviate from within that range. The complete list of parameters is listed in Table 4.4.

Each run lasted for 1,000 "days" as defined in Section 3.2.10. Effectively, this means longer runs (more trades) as the number N of agents increases because of how time is scaled.

4.2.3 Phases

Fig. 4.10 shows sample price series for a strongly negative value of r_p and a strongly positive value indicating a change of character as r_p is varied. For negative r_p the dynamics are dominated by high frequency fluctuations overlaying a relatively small low-frequency component while the reverse seems to be true for positive r_p. This is not entirely surprising because the parameter r_p acts as a kind of autocorrelation between successive price movements. When r_p is negative, an increase in the price will lower the agents' ideal investment fractions, decreasing demand which usually results in a price drop. Conversely, price increases tend to be followed by further increases when r_p is positive because of increased demand.

Hurst exponent

To quantify the dynamics, then, an order parameter which characterizes the autocorrelations is called for. The tickwise autocorrelation (between successive trades) was considered but rejected because it was found to be "noisier" than the alternative—the Hurst exponent. The Hurst parameter $0 \leq H \leq 1$, discussed in Appendix C, quantifies the proportion of high-frequency to low-frequency fluctuations and measures the long-range *memory* of a process. It is an alternative representation of

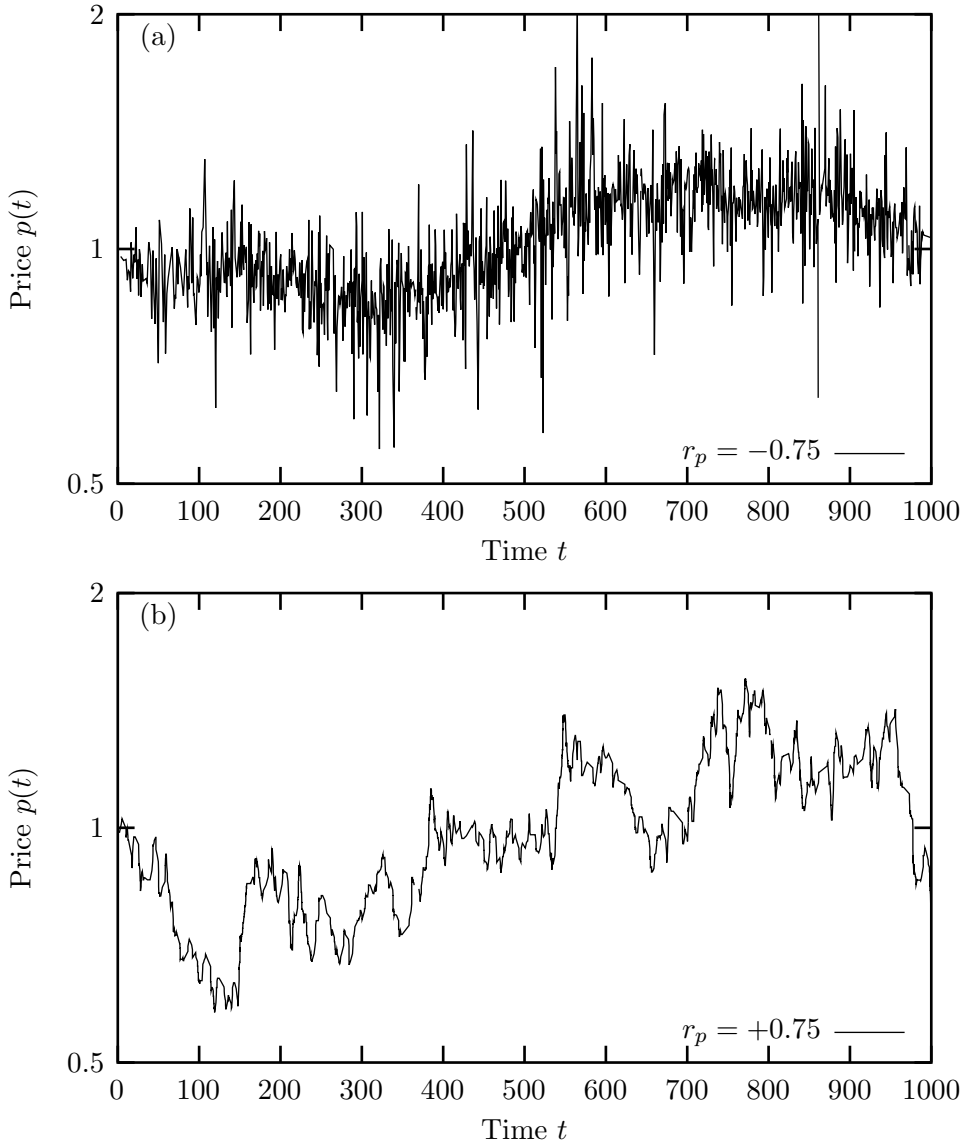

Figure 4.10: Sample price series for DSEM with $N = 100$. Negative values of r_p (a) produce an anticorrelated series while positive values (b) result in positive autocorrelations.

temporal correlations in a time series with $H < 1/2$ for anticorrelated series and $H > 1/2$ for positively correlated data ($H = 1/2$ indicates no correlations).

The Hurst exponent was calculated by the application of dispersional analysis, a simple and accurate method described in Section C.2.1, to the log-return series of the price sampled at discrete intervals of four "minutes" (interpreting a trading "day" to consist of 6.5 hours). Sampling the data at regular intervals did not significantly affect the estimates of H but it fixed the size of the dataset to be analyzed; for large systems many trades could be executed within a few minutes, producing exorbitantly large datasets which were cumbersome to analyze. (Before discretization the largest dataset contained some 400,000 points.) With a fixed sampling rate all system sizes generated the same volume of data, around 100,000 points.

As the price response parameter r_p increases from -0.75 up to 1.00 some interesting properties emerge: for large systems ($N \geq 200$) the Hurst exponent is effectively zero below $r_p \approx -0.4$, indicating very strong anticorrelations in the price series (independent of r_p). Suddenly, near -0.4, the anticorrelations break down and the Hurst exponent climbs quickly (with increasing r_p) to roughly $H \approx 0.4$ (suggesting weakly anticorrelated data). It remains relatively constant until $r_p \approx 1$ where it climbs again, to $H \approx 0.8$.

Interestingly, the Hurst exponent is less than one half at $r_p = 0$ meaning that price movements in one direction tend to be followed by opposite movements even with no explicit price response coded into the agents' behaviour. This arises from corrections to initial over-reaction to news—agents with extremist news responses react to news by placing orders with atypical prices which are corrected for when moderate agents have an opportunity to trade.

Price response greater than unity

Apparently, DSEM exhibits three distinct phases: the first two are demonstrated in Fig. 4.10 but the third ($r_p > 1$) is not shown so it is briefly discussed here. It is characterized by very strong positive correlations in the price series, such that the price explodes or crashes exponentially, depending on the direction of the first price movement.

Very rapidly (within a few "days") the price hits a boundary imposed by a mechanism identical to that in CSEM: the investment fraction in DSEM is actually constrained by δ such that $\delta \leq i \leq 1-\delta$. The price is therefore bounded according to Eqs. 2.43–2.44. This constraint was not mentioned in the development of the model because, in DSEM, the limit is $\delta = 10^{-12}$, a value chosen only to avoid numerical round-off errors. In practice δ was not observed to affect the dynamics whatsoever, except in the region $r_p > 1$.

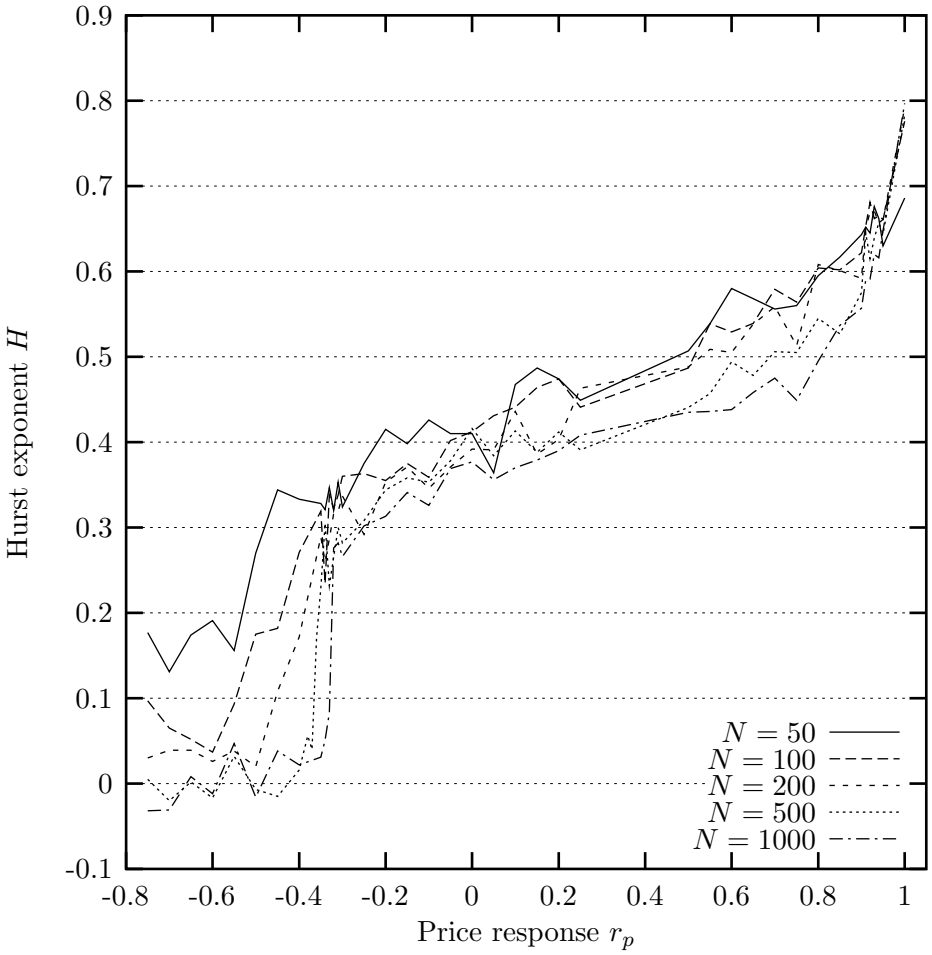

Figure 4.11: The Hurst exponent increases with r_p in DSEM as expected but with two surprising phase transitions emerging at larger system sizes: one near $r_p \approx -0.4$ and the other near $r_p \approx 1$.

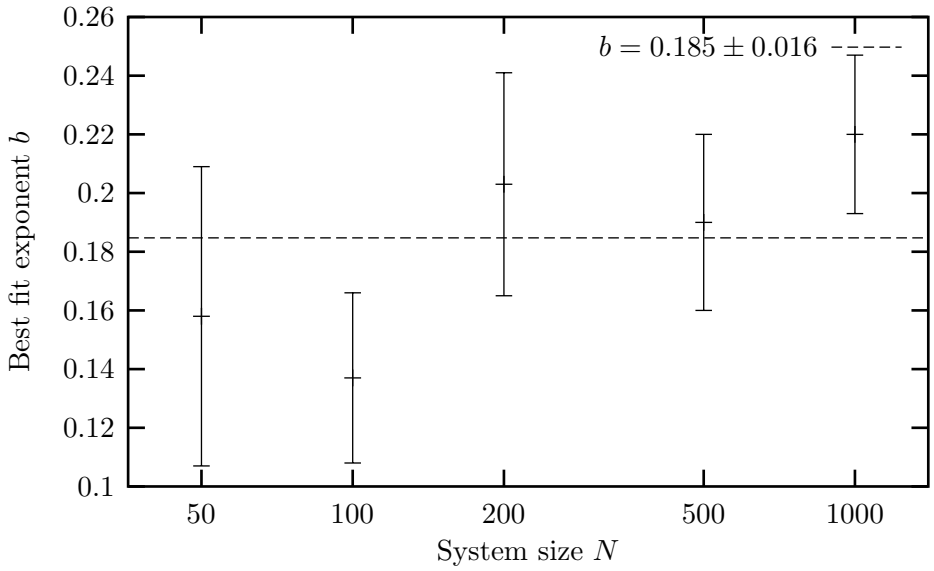

Figure 4.12: The best fits of power laws to DSEM Dataset 1 yield the scaling exponents shown. The average exponent is 0.185 ± 0.016.

The exponential growth of the price in this region indicates that the Hurst exponent (which cannot be accurately measured because the price reaches the boundary too quickly) is identically one $H = 1$, the same value obtained from a straight line (which this is, since we are analyzing the logarithm of the price).

4.2.4 Phase transition to $H = 1$ at $r_p = r_1$

In this section the phase transition for positive values of r_p will be explored. This phase transition is easier to characterize than the one to be discussed in Section 4.2.5 because we have reason to expect the transition to occur at $r_p = r_1 \equiv 1$ and may therefore eliminate one adjustable parameter from the fitting function. (The fit was also performed with r_1 as an adjustable parameter (not shown) and the results corroborate the hypothesis that the transition is at $r_1 = 1$.)

From Fig. 4.11 it is clear that the phase transition at r_1 is not first-order (discontinuous) but appears to be second-order (critical). In the last section it was argued that the transition is to $H = 1$ for $r_p > r_1$ so the power-law to be fit takes the form

$$1 - H(r_p) = C(r_1 - r_p)^b \tag{4.8}$$

with fitting parameters C and b.

The fits were performed over the range $0.75 \leq r_p \leq 0.95$ for each value of N

giving the exponents b shown in Fig. 4.12 with an average value of $b = 0.185 \pm 0.016$. It is not known to what universality class (defined in Section 4.1.7), if any, this transition belongs. Next we explore the phase transition observed for negative values of r_p.

4.2.5 Phase transition to $H = 0$ at $r_p = r_2$

Now we turn our attention to the other phase transition in the system, near $r_p \approx -0.4$. For the smallest systems $N \leq 100$ the transition is not discernable in Fig. 4.11 but it comes into focus as the system size is increased. For intermediate agent numbers, $N = 200$ and 500, the transition looks very much second-order (continuous). However, in the largest system $N = 1000$ the transition is quite abrupt, so particular care must be taken to establish whether it is first- or second-order.

Comments on phase transitions

The distinction between first- and second-order transitions is not merely academic; it can greatly enhance our understanding of the underlying dynamics. First-order phase transitions, characterized by a discontinuity in the order parameter, occur via nucleation: small pockets of the new phase emerge within the old phase and grow until the entire system is in the new phase. On the other hand, second-order transitions, characterized by a continuous order parameter with a diverging derivative, exhibit system-spanning correlations such that the entire system undergoes the transition as a whole.

In the context of the models presented here, correlations would indicate correlated behaviour amongst investors and nucleation would refer to a small sub-group of investors acting differently from the larger population.

Classification of r_2

To explore the phase transition in detail some more data were collected at intermediate system sizes as shown in Table 4.5 for a total of seven different values of N. The datasets were analyzed by attempting to fit both first-order and second-order transitions.

Assuming a first-order transition the fit becomes trivial: we can just assume a linear dependence on r_p in the neighborhood of r_2. (In this case linearity was observed over $r_p \in (-0.3, 0.1)$ for all runs.) The transition point is simply read off the graph from the largest system $N = 1000$ (the transition is resolved with greater accuracy as N increases) giving $r_2 = 0.33 \pm 0.01$. The magnitude of the discontinuity in the Hurst parameter H at r_2 is then $\Delta H(r_2) = 0.281 \pm 0.008$.

Parameters	DSEM Dataset 2	
Particular values		
Number of agents N	300	700
Number of runs	16	16
Common values		
Price response r_p	-0.35 to -0.30 by 0.01 -0.25 to 0.25 by 0.05	
Run length ("days")	500	

Table 4.5: Parameter values for DSEM Dataset 2. These runs are a variation of Dataset 1 (all unspecified parameters are duplicated from Table 4.4) exploring a few other intermediate system sizes.

We now consider the possibility that the transition is second-order. If so, then a power-law dependence

$$H(r_p) = C(r_p - r_2)^b \tag{4.9}$$

should characterize the behaviour near the transition r_2 with adjustable parameters C and b.

Sample fits for $N = 500$ and 1000 are shown in Fig. 4.13 with a simple linear fit representing a first-order transition for comparison. Notice that the $N = 500$ system is better described by a critical transition but N increases to 1000 the transition becomes sharper, more like a first-order transition suggesting that the scaling behaviour is only a finite-size effect.

Finite-size scaling

Assuming criticality, each system in Datasets 1 and 2 were fit to Eq. 4.9 and the best-fit exponents b are plotted in Fig. 4.14. Clearly, the exponents exhibit a trend as the system size N increases. On a log-log graph the trend appears linear suggesting that the exponents scale with the system size as yet another power law $b \propto N^m$. (Be warned, this is not a traditional—nor rigorous—finite-size scaling argument.) The scaling exponent is found to be $m = -0.25 \pm 0.04$ meaning that the exponent b will be halved every time the system size is scaled up by a factor of 16 and in the thermodynamic limit ($N \to \infty$) b drops to zero.

To understand what is going on here, consider the general scaling function

$$y = c(x - x_c)^b \tag{4.10}$$

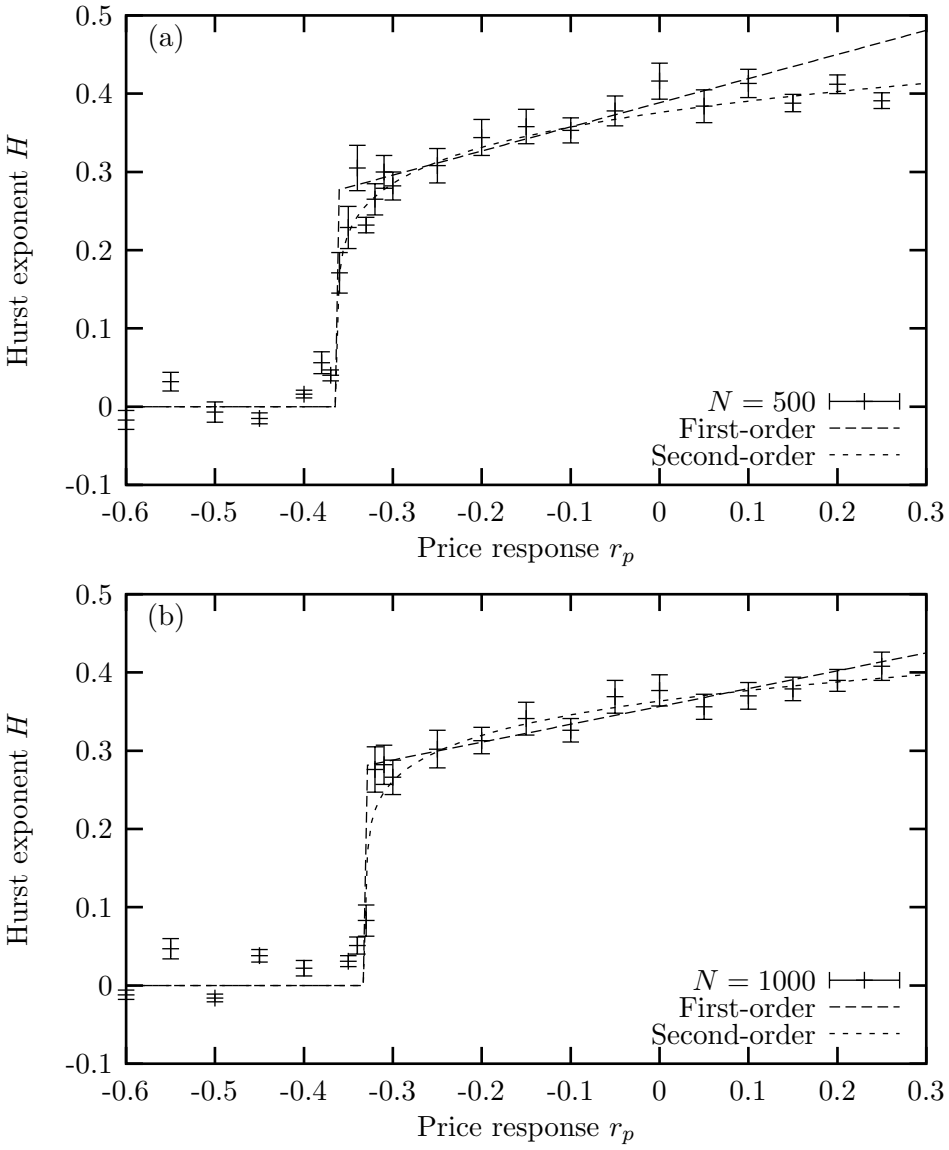

Figure 4.13: Sample fits of first- and second-order phase transitions to $N = 500$ (a) and $N = 1000$ (b) near r_2 in DSEM show that the power-law fits better for small N but the first-order prevails for larger systems.

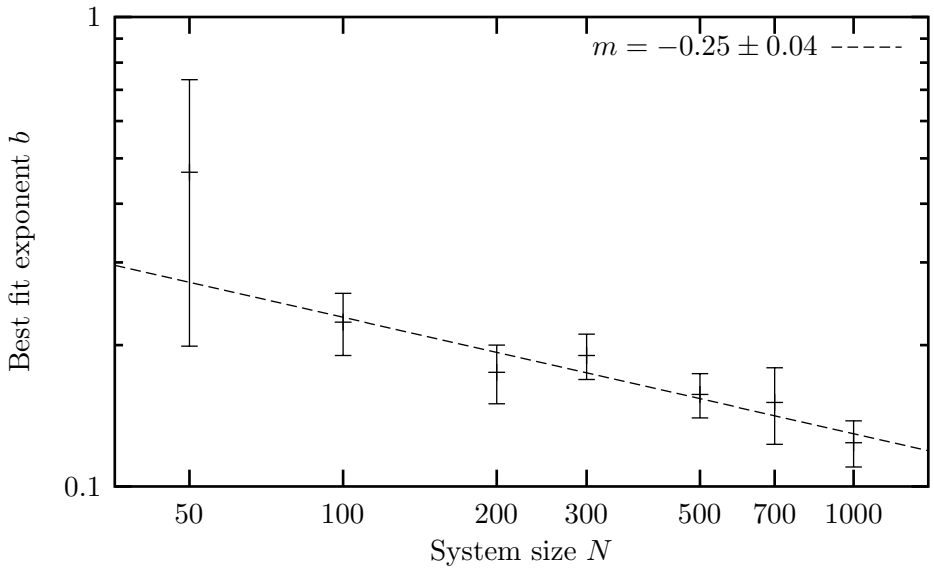

Figure 4.14: As the system size N increases the critical exponent b tends to zero. The line represents a power-law fit $b \propto N^m$ giving an exponent $m = -0.25 \pm 0.04$.

which has a slope

$$y' = bc(x - x_c)^{b-1}. \tag{4.11}$$

If we demand that the scaling function mimic a first-order transition, requiring that both $y > 0$ and y' approach constants as $x \to x_c$ then the scaling exponent must vary such that

$$b = \frac{y'}{y}(x - x_c). \tag{4.12}$$

So a second-order transition "mimics" a first-order in the limit $b \to 0$.

Returning to the transition at $r_p = r_2$ in DSEM we see that the apparent criticality is an artifact of finite simulation size and in the limit $N \to \infty$ the transition is first-order.

Intermittency

As discussed above, there are important consequences of knowing a transition is first-order. The foremost is that fluctuations are local, they do not spread throughout the entire system (as is found for critical points). Another consequence is that the transition is a change of *quality* not simply of quantity. That is, since the order parameter exhibits a discontinuity the nature of the system changes qualitatively, not just quantitatively. A third important feature of first-order transitions is *nucleation*. Near the transition stochastic fluctuations can often give rise to small pockets

which exhibit one phase while the greater system is within the other phase. A good analogy to keep in mind is a pot of boiling water: steam bubbles form on the bottom and sides of the pot where small variations of the surface exist.

The last interesting property of first-order transitions that will be mentioned here is intermittency. As discussed above, near the phase transition bubbles of one phase form at nucleation points within the other phase. In a spatially-extended system this causes intermittent periods of either phase at any particular point in the system. Since DSEM is nonspatial the intermittent behaviour is captured in the price series which is observed to consist of periods of low activity separated by periods of high activity, as demonstrated in Fig. 4.15.

4.2.6 Summary

In this section we explored the phase space of the price response parameter r_p and the number of agents N in DSEM. Two phase transitions were observed: at $r_p = r_1 = 1$ DSEM undergoes a critical transition to perfectly correlated price movements (the Hurst parameter H goes to unity), and at $r_p = r_2 \approx -0.33$ a first-order transition is observed. Below the transition strong anticorrelations in the price series are observed but above it only weak anticorrelations exist. Near the transition point the system spends time in both regimes giving rise to clusters of high volatility.

4.3 Number of investors

Before comparing the models with empirical data we should complete our exploration of the phase space. In both CSEM and DSEM we explored a variety of system sizes (number of agents N) in order to enhance the resolution of the phase transitions. In this section we will re-evaluate this data in light of the discovery that in many market simulations the dynamics reduce to being semi-regular in the limit of many investors [62, 63].

Most past simulations were performed with investors numbering between 25 and 1,000 [28–30, 32, 33, 47] with the largest ranging between 5,000 and 40,000 [17, 34, 35]. Even the largest of these are minuscule when compared with natural systems exhibiting phase transitions, which have on the order of 10^{23} particles. It is not clear that these market models (including CSEM and DSEM) are at all interesting in the limit of many investors. In fact it has been discovered that the dynamics of many of these models become almost periodic as the number of investors grows [62, 63].

Whether this behaviour detracts from the models is uncertain. The models can only be tested by comparison with real markets but even the largest markets

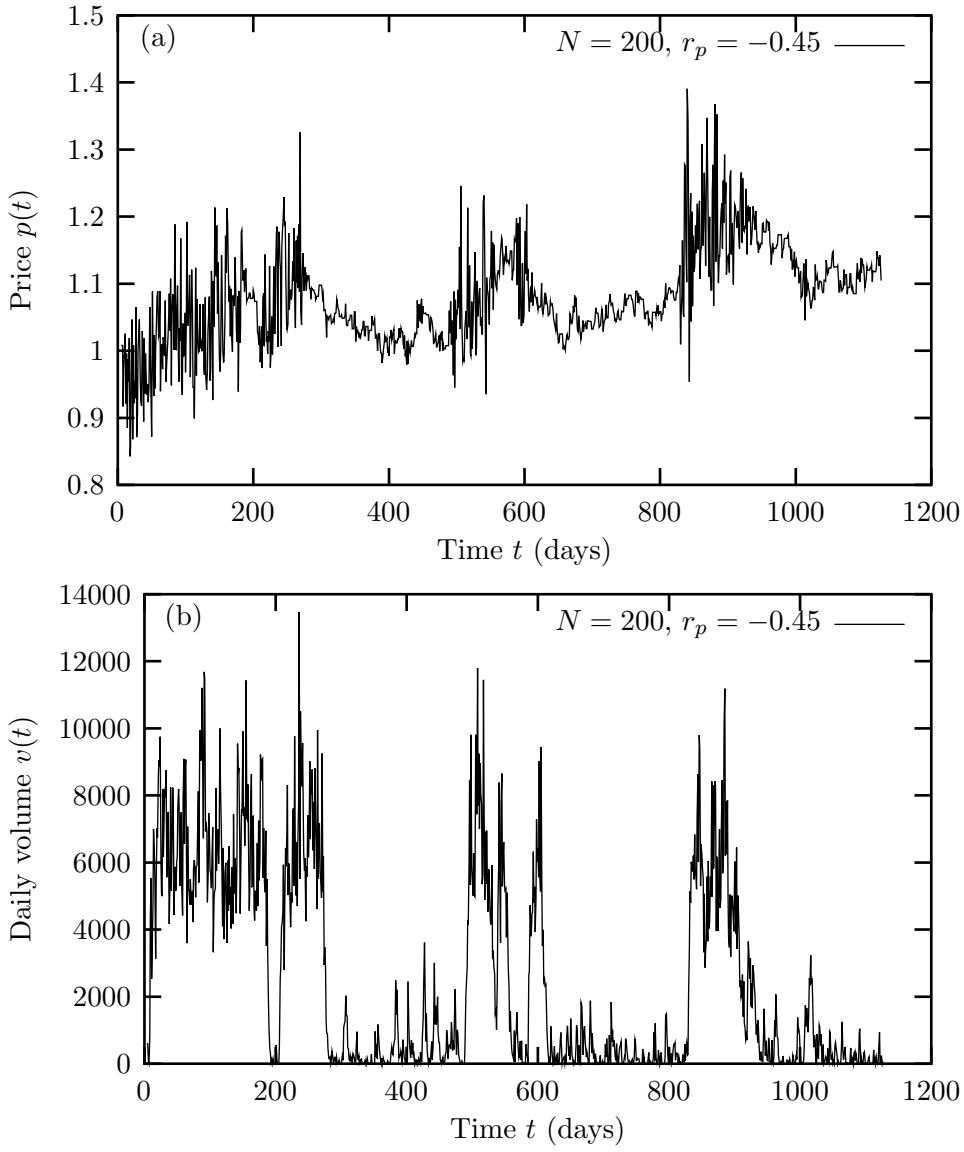

Figure 4.15: The price series (a) and daily volume (b) of DSEM with $N = 200$ and $r_p = -0.45$ is a good example of intermittency. The dynamics fluctuate between two phases.

Parameters	CSEM Dataset 3		
Number of agents N	10,000	10,000	10,000
Forecast error σ_ϵ	0.01	0.08	0.15
Investment limit δ	10^{-2}	10^{-2}	10^{-2}
Number of runs	1	1	1
Run length (time steps)	10,000	10,000	10,000

Table 4.6: Parameter values for CSEM Dataset 3. These runs are a variation of Dataset 1 (all unspecified parameters are duplicated from Table 4.1) with many agents $N = 10,000$.

cater to an infinitesimal number of individuals when compared with natural systems. After all, the natural world consists of only a few billion "agents" of which only a minute fraction are actively involved in stock trading. Therefore, one may argue that these models do not need to exhibit rich behaviour in the limit of many investors in order to be realistic. They need only exhibit realistic dynamics on the same scale as real markets.

Nevertheless, it is interesting and useful to understand how the models predict the dynamics to evolve with increasing investor numbers. One advantage is that testable predictions may be made with regard to how a market will scale as more investors come aboard. In the last few years the number of investors in the markets have grown substantially, mainly due to the rise of the internet which allows traders to monitor their portfolios in (almost) real time and execute trades promptly. The only research the author is aware of to explore the consequences of growing markets is a model which suggests that fluctuations increase with system size [18].

Thus, it is useful to explore the effect of increasing the number of agents in both the Centralized and Decentralized models.

4.3.1 Centralized Stock Exchange Model

To test the effect of changing the number of agents N thoroughly a new dataset (see Table 4.6) was collected with markets containing $N = 10,000$ agents in three regimes: far below the critical point $\sigma_\epsilon = 0.01$, near the critical point $\sigma_\epsilon = 0.08$, and above the critical point $\sigma_\epsilon = 0.15$. The price series for the first and last of these are shown in Fig. 4.16 and 4.17, respectively.

Far below the critical point the dynamics appear to be largely invariant under change of the number of investors. Interestingly, the dynamics do display semi-periodic intervals, as demonstrated in Fig. 4.16(b), but this occurs for all system

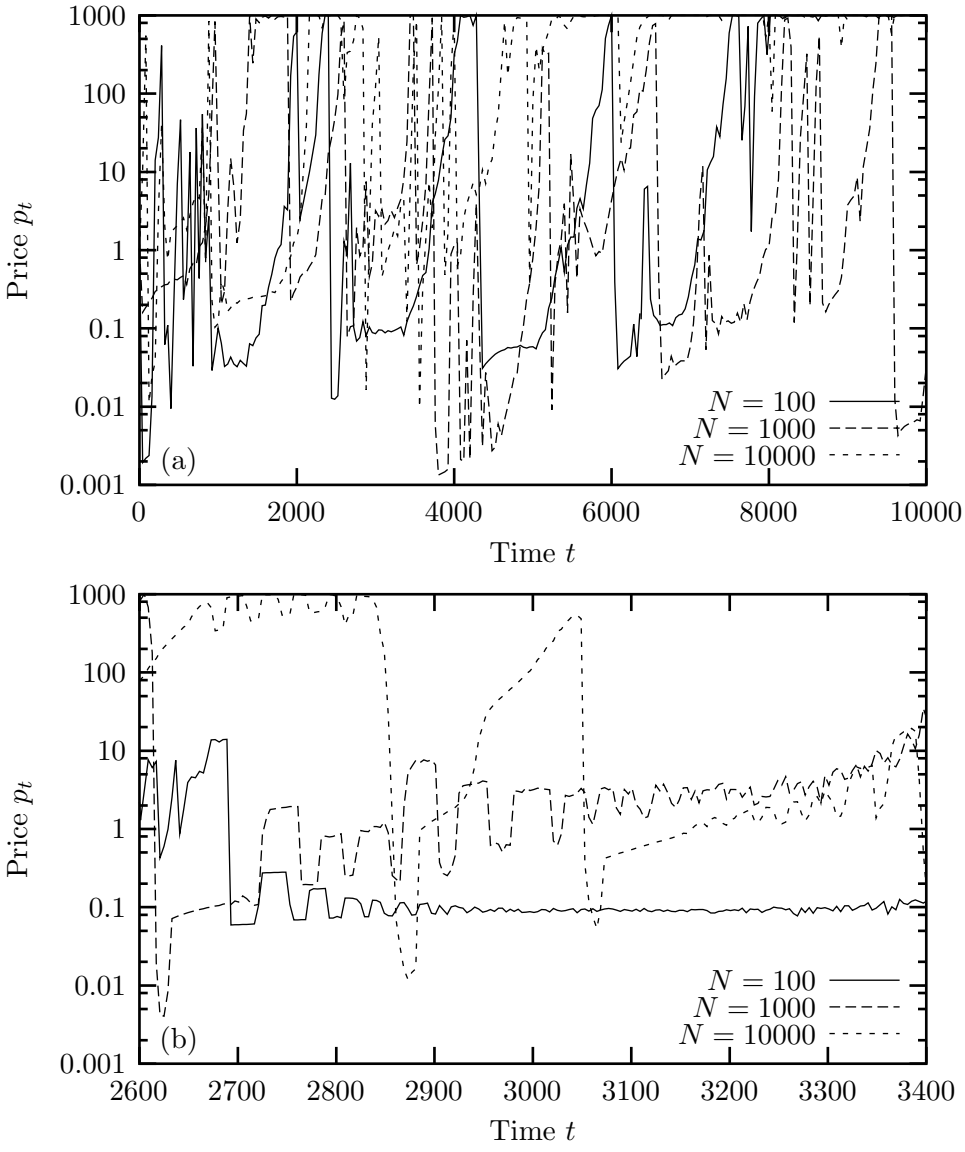

Figure 4.16: The price series of CSEM for $\sigma_\epsilon = 0.01$ (a) appears unaffected by changing the number of agents N. In particular, occasional semi-periodic fluctuations (b) are observed for all system sizes.

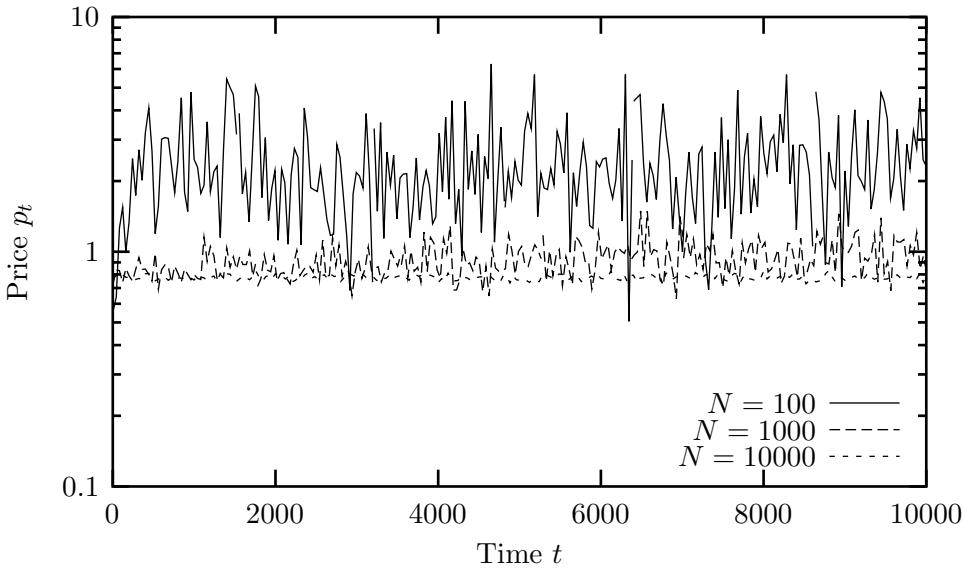

Figure 4.17: The price series of CSEM for $\sigma_\epsilon = 0.15$ exhibits smaller fluctuations and a lower mean as the system size increases. (The lower mean may simply be because the system has not reached a steady state yet.)

sizes and is not a stable phenomenon even for the largest system so it does not appear that the dynamics converge to being almost periodic in the limit $N \to \infty$ as found in other models [62, 63].

Above the critical point the most obvious feature is that the fluctuations decline with the number of investors (see Fig. 4.17). But this is to be expected since the trading price results from the interactions between N "noisy" investors. Therefore we expect the fluctuations (measured as the standard deviation of the log-price) to decrease as $1/\sqrt{N}$, a hypothesis which was found to hold fairly well for $\sigma_\epsilon = 0.15$ between $N = 50$ and $N = 10,000$ but which held better further from the critical point (as can be seen in Fig. 4.7 which demonstrates the variances multiplied by system sizes nearly collapse to a single curve).

The main conclusion to be drawn from this analysis is that the dynamics of CSEM do not appear to become trivial or converge to a semi-periodic pattern as the number of investors becomes infinite. Above the critical point fluctuations do diminish but as one approaches the critical point they reappear and below the critical point the dynamics are largely independent of the number of agents.

4.3.2 Decentralized Stock Exchange Model

Collecting data for $N = 10,000$ in DSEM proved prohibitive since each run would require a full week of computer run-time to collect sufficient data for analysis. This occurs because DSEM requires on the order of N^2 operations per simulation "day" (N calls per day with N potential replies each) whereas CSEM grows linearly with N. Therefore the data collected in Table 4.4 will be used here.

Again, we observe the price series for a variety of system sizes—this time $N = 50$, 200, and 1000 agents—and regions of phase space spanning the phase transitions—$r_p = -0.75$, 0.00, and 0.90—to estimate how the dynamics would change as the system size grew without limit. In each region it was found that the fluctuations actually grew with system size, especially below the first-order transition $r_p < r_2 \approx -0.33$.

The two plots in Fig. 4.18 show the price series only for $N = 50$ and $N = 1000$ ($N = 200$ exhibited predictably intermediate fluctuations and was not plotted to reduce clutter). The plots show that the fluctuations are somewhat larger for the larger system when $r_p = 0.90$ and significantly so for $r_p = -0.75$. Thus the dynamics do not reduce to semi-regular in the limit of many investors.

4.3.3 Summary

In this section we tested the effect of increasing the number of agents in both CSEM and DSEM in light of recent research that indicates that some market models become quasi-periodic in the limit of many agents [62,63]. It was found that fluctuations did decline in CSEM when the forecast error σ_ϵ was significantly above its critical value but the dynamics were largely invariant under variation of the number of agents below the critical point.

In DSEM the fluctuations actually grew as more investors were introduced, especially below the first-order transition at a price response of $r_p \approx -0.33$. This makes an interesting and testable prediction regarding empirical markets: it indicates that fluctuations in stock prices should be greater in larger markets. (Some caveats are required: the size of the market is measured in terms of the number of independent investors (a fund group would be interpreted as a single investor) and not the total value of the outstanding stock. Recall that in these simulations the total cash and shares were held fixed: with more investors each held a smaller portion of the total resources.)

Since the number of investors does not strongly affect the dynamics we are free to choose values which correspond well with observed market fluctuations. Comparing the fluctuations with empirical data will be the subject of the next chapter.

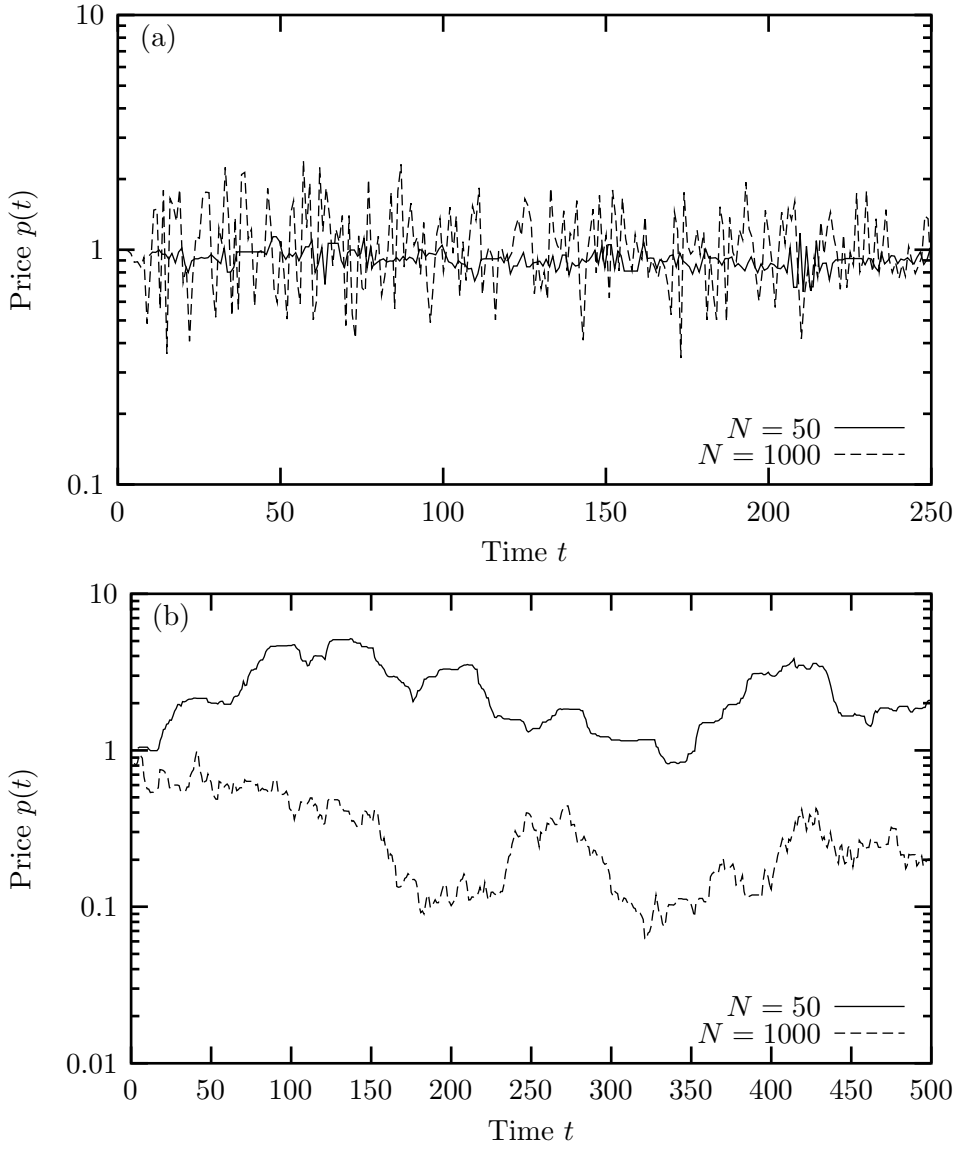

Figure 4.18: In DSEM the price series does not get more regular as the system size is increased—in fact the fluctuation grow. This is especially true for $r_p = -0.75$ (a) but it is also indicated to a lesser degree at $r_p = 0.90$ (b).