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 1

Introduction

1.1 Financial markets

Financial markets include *stock markets*, such as the New York Stock Exchange (NYSE), which deal in ownership *shares* of publicly-owned companies. Companies owned privately can raise equity capital through an *initial public offering* (IPO) which releases part-ownership to the public. When the public *stockholders* wish to sell some of their shares they do so on a financial market. The market typically charges the company a fee to list it and may require the company to meet certain standards in order to protect investors.

On the NYSE, trades are handled by a restricted number of *brokers* who are governed by the market's rules and regulations. Brokers receive trading orders—consisting of a quantity of shares to trade and (optionally) a price—from the public and bring them to *specialists* who deal only with specific stocks.

The specialist's role is to compare the highest *bid* (buy order) price with the lowest *offer* (sell order) price and if they meet, execute the trade. At the beginning of trading each day the specialist also finds a fair market price for a stock by balancing the outstanding *supply* (total offers) with *demand* (total bids). Although actually more complicated, for our purpose this is a sufficient explanation of the specialist's role.

1.2 Motivation for research

Neglecting *dividends* a company may pay to its shareholders, investors make money on the market by buying stocks at low prices and selling them at higher prices. But given identical (publicly available) information one would expect (similar) investors to have the same prediction for how the price would move and they should all place similar orders.

For example, if Betty hears that company XYZ has discovered oil, she may well expect the company's stock price to climb and so she places a bid order. The trade will not be exercised, however, until another investor, say Sam, offers to sell his shares in XYZ. But the question is then raised in Betty's mind, "Why does Sam want to sell?" Is he being irrational? Did he miss the good news? Does he know something Betty doesn't?

Conventional wisdom assumes Betty and Sam have slightly different expectations about the future price of XYZ, perhaps due to imperfect information. Thus the market dynamics are driven by random fluctuations. But this assumption leads to predictions that price fluctuations should be normally distributed (or log-normally) and that trading volume should be low and steady.

What is actually observed on all markets is bursts of activity with very high volume and/or extremely large fluctuations in price which occur much more frequently than the conventional wisdom can account for. The reason for these bursts has not been established and is an interesting topic of research.

1.2.1 Motivation for the physicist

Statistical physicists and condensed matter theorists have developed a significant arsenal of tools for analyzing many-particle systems with strong, localized interactions. Methods such as mean-field theory, the renormalization group, and finite-scaling analysis allow physicists to explore complex, irreducible systems such as spin glasses (highly disordered magnetic systems) where the important details are in the interactions between the particles, rather than the individual particles themselves.

Recently, physicists have realized that the methods developed above may also be useful for non-physical systems such as ecological and social systems. The leap of faith required is the assumption that it is not necessary to fully understand the individuals in the system themselves (their motivations, for instance), but only to the point that one can construct reasonable rules for the interactions between individuals.

Whether this leap of faith is justified remains an open question but interest is mounting within the physics community in complex, socio-economic systems like the stock market. In 1995 the Los Alamos National Laboratory (LANL) condensed-matter preprint archive (`http://arXiv.org`) accepted three papers containing the word "market" in their abstracts (a check was done confirming they were all finance-related) representing 0.16% of the submissions that year. Since then it has doubled every year through 1999 when fifty of the 5,490 (0.91%) submissions were market-related. (The foray of physicists into economics has come to be known as *Econo-*

physics.)

Phase transitions

Together with the analytic tools physicists bring to the subject, they also bring a fresh perspective and new questions. For instance, it has been empirically observed that price fluctuations exhibit *scaling* [4–7], meaning the fluctuations appear invariant under a change of scale, over orders of magnitude from a few minutes to a few days. Scaling is characterized by power-law distributions which are very familiar to physicists because they occur near second-order or *critical* phase transitions.

Phase transitions, in the context of thermodynamics, are well understood phenomena. The terminology of *order parameter*, a dependent variable which undergoes a "sharp" change, and *control parameter*, the variable which is smoothly adjusted to produce the change, is used to quantify the transition. In the case of *first-order transitions* (such as melting) the order parameter undergoes a discontinuity—it jumps to a new value. The jump is accompanied by an absorption or liberation of energy (latent heat). Usually, fluctuations within the substance can be ignored for first-order transitions.

To demystify the above definitions, consider a pot of water boiling at 1 atmosphere of pressure and 100°C. If we choose temperature as the control parameter then density could play the role of the order parameter. Below 100°C water is a liquid with a relatively high density. Above this point, all the water is in the form of steam which has a significantly lower density. At the transition we observe fluctuations in the form of small, uniform steam bubbles. This is a first-order transition.

In contrast, second-order, or *critical*, transitions are characterized by a discontinuity in the derivative of the order parameter (see Fig. 1.1). In fact, the derivative diverges at the critical point. Further, near the transition, properties are dominated by internal fluctuations on all scales. For example, let us revisit our pot of boiling water. We raise the pressure to 218 atm and the temperature to 374°C (the critical point of water). Again, we observe steam bubbles but in this case the bubbles exist on all scales—from microscopic to the size of the pot itself [8]. Also, the density (but not its derivative) is continuous across the transition.

Near a critical point, many thermodynamic properties obey diverging power laws. Early studies of critical phenomena revealed that the characteristic exponents for the power laws clustered around distinct values for a variety of systems. This suggested that some of the features of separate systems were irrelevant—they belonged to the same *universality class.* Some of the irrelevant variables in a universality class are usually the type of local interactions, the number of nearest neighbours, et cetera. On the other hand, dimensionality and symmetry, for example, appear to

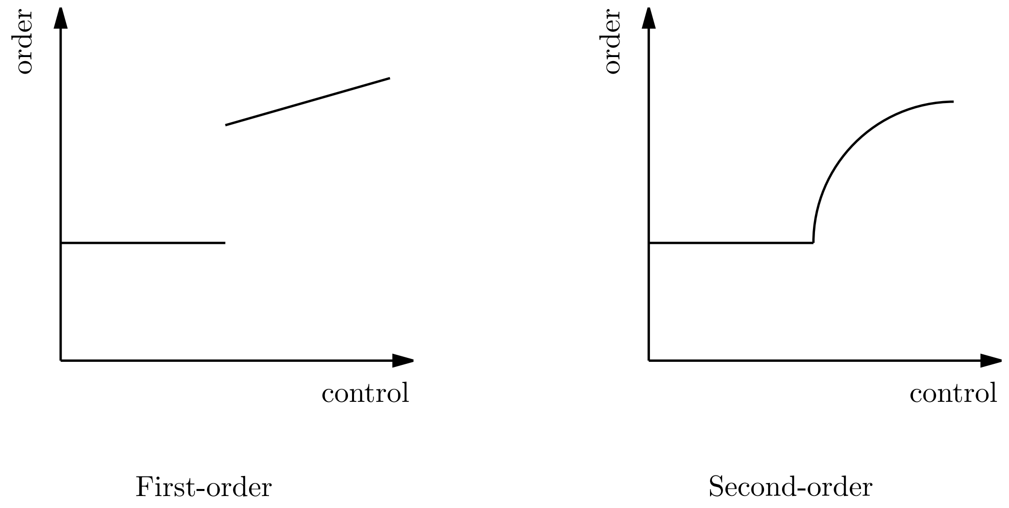

Figure 1.1: Sample phase transitions. A first-order transition (left) is characterized by a discontinuity in the order parameter, while a second-order (critical, right) is discontinuous in the first derivative.

be relevant variables—that is, change the number of dimensions or the symmetry laws and the system changes its class (or may even cease being critical). Critical systems with the same relevant variables but different irrelevant variables have the same critical exponents and are said to belong to the same universality class.

Returning to our discussion of markets, the estimated power law exponent in the financial data seems to exhibit universality. That is, the exponents seem to be similar for a number of different markets and stocks and they also seem not to change over time [6, 7, 9, 10]. This evidence suggests that markets operate at or near a dynamical critical point as studied by physicists.

Self-organized criticality

To a physicist, the question of whether the market operates at a critical point is especially interesting. The traditional theory of critical phenomena states that a system will approach a critical point via deliberate tuning of the control parameter. In the above example, by adjusting both the temperature and pressure, water was brought to its critical point.

This description does not seem to apply to markets, however. The rules governing market dynamics were not chosen in order to put the market in a critical state. In fact, there does not appear to be any analog for temperature, which could be used to explain why the market might be at a critical point. If it is critical, it appears to have arrived there spontaneously, without any tuning of a control parameter. This phenomenon has come to be known as *self-organized criticality* (SOC) and was originally proposed as a possible explanation for scaling in many natural phenomena [11, 12].

The canonical example of SOC is a pile of sand to which grains are added very slowly. As each grain is dropped it may cause the local slope of the pile to exceed a threshold and collapse, dispersing grains within a local neighbourhood. These grains may cause further instabilities producing a cascade reaction. Measuring the total effect of dropping each grain yields a power-law distribution of avalanche sizes, indicating the presence of a critical point. The criticality is said to be self-organizing because it emerges spontaneously from the simple process of dropping grains periodically.

In some cases, SOC can be mapped back onto traditional criticality by a separation of timescales: systems which responds quickly to very slow driving forces are candidates for SOC. In particular, the sand pile model described above qualifies for this mapping because the driving force (dropping of grains) is much slower than the duration of the avalanches [13].

More generally, the appearance of SOC can be an artifact of how the system is constructed. Some natural choices of parameter values (such as an infinitesimal driving rate, as discussed above) automatically lead to dynamics which can be critical or very nearly so. Traditional criticality is only revealed when the parameter is manually varied [14] (for example, by increasing the rate at which sand is added to the pile).

Whether the markets operate at a critical point and, if so, how they develop towards and maintain this state is of interest to physicists.

1.3 Anticipated challenges

Although synthetic constructs, the markets are difficult to study scientifically for many of the same reasons as natural phenomena. Firstly, stocks are strongly coupled to each other and to other systems, both natural and man-made. For example, an earthquake in Taiwan on September 20, 1999 which cut off electrical power at Taiwan Semiconductor Manufacturing (TSM) and significantly disrupted production, had only a minimal impact on the company's stock price. However, their South Korean competitors' stock prices soared in anticipation of increased demand.

This highlights the second challenge in studying the market: investors' responses (and hence stock price fluctuations) to incoming news can be strongly nonlinear. A company's quarterly forecast of a loss of ten cents per share could conceivably have much more than double the impact of five cents. The precise response function (if one exists) is unknown.

Thirdly, the impact of an exogenous event may be practically, or even theoretically, unquantifiable. Investors may receive imperfect information and/or the necessary calculations to assess the impact of the news on a stock's price may be too complex, beyond the rational abilities of the investor.

Lastly, only some of the information which drives investors' actions is broadcast to all. The rest (a rumour, for instance) is transmitted through a complex network of friends, families, and co-workers. It is not clear if this information can be neglected and, if not, how the network is to be represented, structurally.

1.4 Modeling

The natural sciences are well acquainted with these types of challenges and their reaction is to study the system in two ways: first empirically, then with an idealized representation.

Empirical analysis is the first and best way to understand the world around us. By collecting data and studying statistical properties thereof we can learn about the underlying distributions governing many phenomena. Then, once sufficient empirical data have been collected idealized models may be constructed to try and account for the data.

A vast store of financial market data is available. For instance, precious metal price data are on record all the way back to the 1200s [15]. A large number of these data sets have been analyzed and the results indicate that large market fluctuations (outliers) occur much more frequently than would be expected (the frequency distributions exhibit *fat tails*) and, unexpectedly, fluctuations occur in clustered bursts of volatility rather than uniformly. This thesis will not focus on empirical analysis of financial data, rather relying (mainly) on these published results.

Instead, this thesis will focus on idealized representation or modeling of financial markets. Social scientists have developed simple analytic models of the stock market. For tractability they assume a small number of investors who have *perfect rationality* (unlimited computational power) and complete information [16–18]. These models are interesting to economists because they can explain equilibrium stock prices [19]. However, they are uninteresting to the physicists for precisely the same reason—since they are *equilibrium* models they fail to exhibit fat-tailed fluctuation distributions or clustered volatility.

1.4.1 Computer simulations

My hypothesis is that the complex dynamical behaviour of the stock market is an emergent property arising from the interactions of many agents and is largely independent of the complexity of the agents themselves. In order to test my hypothesis I will construct some simple models meant to capture the essence of the stock market and study them experimentally via computer simulation.

Computer simulation is necessary because many-agent models are impractical or even impossible to analyze by hand—the number of interactions which need to be accounted for typically grows as the square of the number of agents. (Even the simplest models tend to be too complex for an analytic treatment.) So we turn to computers, which are capable of performing millions of calculations rapidly, with no (significant) errors.

There are many objections to working with computer simulations but some of these apparent shortcomings are actually advantages. For instance, it is impossible to construct a many-agent simulation with perfect rationality and complete information: each agent's expectations are formed on the basis of every other agent's expectations, which are formed on the basis of every other agent's expectations, ad infinitum. (In some cases this infinite regress can be collapsed and solved.) Besides being impossible to incorporate into a simulation, I hope the reader will agree this is an unrealistic account of investor behaviour.

Simulation may also seem inappropriate because, to develop a *stochastic* model, random events must be incorporated but computers are incapable of generating truly random numbers. As an alternative, a number of algorithms have been constructed to produce *pseudo*-random numbers which pass all known statistical tests for randomness [20–22]. However, these generators still require a *seed* from the user—a random, initial number to begin the sequence. This flaw can often be a blessing because it offers *replicability* in one's experiments—by seeding the simulation with the exact same number as a previous iteration, the entire time series can be reproduced. (To generate independent time series different seeds are used.)

Finally, market micro-simulation may be objected to because the events (for example, news releases) which drive the dynamics must be explicitly coded into the simulation. As discussed above, these events are often not even quantifiable and, hence, can not be accurately coded. However, turning this argument around, this is yet another advantage of the simulation methodology. Any number of alternate hypotheses of the structure of the driving events can be encoded and their impact on the dynamics tested experimentally. One of the interesting questions this thesis will address is "How complex does the input (news) need to be to produce realistic output (price fluctuations)?"

1.4.2 An appeal for simplicity

A common temptation when constructing computer simulations is to try to capture as much detail as possible in order to make the simulation realistic. But there are a number of reasons the model should be kept simple: Firstly, model complexity must be balanced against the constraints of current computational speed. Simple models require less computational power and produce larger datasets. Since large quantities of data are required to test the frequency distributions of rare events (such as price crashes), simpler is better for our purposes.

Secondly, as a model's complexity grows its capacity for being understood diminishes. Some global climate models (GCMs), for example, have reached sufficient complexity that the modelers specialize in only a particular subroutine of the model, such as cloud formation. Very few (if any) of the researchers have a full grasp of every detail of these simulations. A problem with this approach is that the model becomes as difficult to understand as the system it was meant to idealize—a problem known as Bonini's Paradox [23]. (Of course, GCMs are extremely useful for predictive purposes, but perhaps not for furthering scientific understanding.)

Thirdly, by starting with a trivial model and gradually adding layers of complexity, it is possible to determine the minimum requirements for a model which captures the essence of the system under investigation. In the case of the market model this could mean building on a simple model until fat-tailed distributions (for example) are observed in the price fluctuations. Then we can say with some confidence, "These ingredients are the minimum requirements to explain market fluctuations."

Finally, there is the issue of *Occam's razor*. In the 1300s the Franciscan monk, William of Occam stated, "Causes are not to be multiplied beyond necessity" [24,25] or, to paraphrase, "The simplest explanation is best," guiding the course of science for centuries. Notice this claim is aesthetic, not epistemological—it does not claim that the simplest explanation is *true*, but simply to be preferred, at least until evidence comes to light which requires us to reject it. In Bayesian probability theory, Occam's razor has an even more precise role: given two theories which explain a phenomenon equally well, the one with fewer adjustable parameters is assigned a greater numerical likelihood [25, Ch. 24]. Similarly, we should construct models which contain as few parameters, or assumptions, as possible.

1.5 Organization of the thesis

In this thesis I will develop and implement via simulation two hypothetical models of stock exchange. An early model, which introduces the idea of a centralized market, will be described in Chapter 2, and a later model, which discards the centralized trading restriction, in Chapter 3. In Chapter 4 the phase space of these models will be explored revealing some interesting phase transitions, including a critical point in either model. Then, in Chapter 5 experiments will be performed and the results of the two models will be compared with each other and empirical data. It will be discovered that the centralized model is incapable of generating the desired dynamics but the decentralized model can exhibit both fat tails and clustered volatility.

Some interesting results of an experiment in investing, using a hypothetical portfolio, will be discussed in Chapter 6. The thesis will close with a discussion of some conclusions which can be drawn from the research and some ideas for future research.