B.A.J. **2**, IV, 809-973 (1996)

FINANCIAL ECONOMICS — AN INVESTMENT ACTUARY'S VIEWPOINT

BY R. S. CLARKSON, B.Sc., F.F.A.

[Presented to the Faculty of Actuaries, 18 March 1996]

ABSTRACT

The paper compares and contrasts the theories and methodologies of financial economics with generalisations drawn from practical experience of institutional investment management. A Dynamic Equilibrium Model is put forward as a much more realistic theoretical framework than the Sharpe Diagonal Model, a downside theory of risk is developed from sports analogies, and a detailed model of human behaviour, incorporating benchmark levels of 'unintelligent', 'intelligent', 'optimal', and 'rational' behaviour, is suggested as a replacement for the very restrictive simplifying assumptions of existing theories. On the basis of a classification system derived from parallels with the history of astronomy, it is concluded that, as a scientific framework, the current theories of financial economics are seriously incomplete, the two crucial areas of incompleteness being the absence of a downside theory of risk and the absence of a realistic model of human behaviour under conditions of risk and uncertainty.

KEYWORDS

Financial Economics; Investment; Risk; Uncertainty; Rational Behaviour; Modern Portfolio Theory; Efficient Market Hypothesis; Utility Theory; Stochastic Models

1. INTRODUCTION

1.1 *Objective*

The objective of this paper is to put forward for discussion what I see, after a quarter of a century's practical experience of institutional investment, as a better conceptual framework for the theory of finance and investment than that provided by the body of theory and empirical research generally known as financial economics.

1.2 *The Actuarial Approach to Investment*

1.2.1 Crucial actuarial skills are to be able, first of all, to identify real-world data that are relevant to a practical business situation, and then to process these data into an appropriate numerical framework to which human judgement can be applied. Many of the most important investment papers in the United Kingdom actuarial literature take a very simple (but nevertheless very perceptive) idea, and then apply precisely these actuarial skills to develop it into a powerful practical tool.

1.2.2 Two excellent examples of this approach are Pepper (1964) and Weaver & Hall (1967). In the former case, the basic idea is that the deviation of the gross redemption yield of a gilt-edged stock from a yield curve fitted to all stocks in a particular maturity range will be a very powerful indicator of that stock's cheapness or dearness relative to the market. In the latter case, the basic idea is that a linear regression analysis which compares the ratings of ordinary shares against some 'best fit' using relevant fundamental attributes, such as expected earnings growth rates and dividend payout ratios, will be a very useful means of identifying shares that are likely to either outperform or underperform the market over the medium term.

1.3 *Practical Relevance and Mathematical Elegance*

When actuaries build investment models, their over-riding concern, from a mathematical point of view, is to put in place a robust computational framework that will work day in, day out, with minimal diagnostic checking being required to ensure that nothing untoward is happening behind the scenes. Instances of mathematical elegance, or even beauty, may arise from time to time, but such features are of no more than passing interest to the actuary. In Clarkson (1978), for example, most of the mathematical content relates to the iterative procedures required to obtain a practical solution to the elegant, but initially quite intractable, transcendental equation that I obtain for the expected price of a gilt-edged stock.

1.4 *The Rise of Modern Portfolio Theory*

There are, I believe, three main reasons why the general methodology which became known as Modern Portfolio Theory achieved such rapid acceptance in many quarters around 1970. First, there was, previously, no self-contained body of finance theory that could be taught in business schools. The Sharpe Diagonal Model, which is linear in nature, and uses the 'beta' of a security as a measure of risk, was an apparently plausible investment model which very quickly filled the previous vacuum. Second, the axiomatic treatment that became possible, once risk was equated to variance (or, equivalently, standard deviation), transformed the investment field from a practical discipline, guided only by a few vague principles, into a new and exciting branch of applied mathematics. This mathematical elegance, as exemplified by influential textbooks such as Sharpe (1970), attracted vast numbers of new entrants to the field, virtually all of whom were prepared to accept, at face value, the apparently innocuous simplifying assumptions that were necessary to make the new approach possible. Third, the new approach appeared to give precise numerical answers in areas where the 'traditional' approach was perceived to be unable to give anything more than vague qualitative guidance. The epithet 'quantitative' soon came into general usage, and this was interpreted in many quarters as meaning 'more scientific', thereby giving greatly enhanced credibility to the new approach.

1.5 *Different Scientific Worlds*

1.5.1 For the past twenty years or so, I have lived in a different scientific world from the proponents of Modern Portfolio Theory, or what would now be called financial economics. This is not because I cannot grasp the underlying mathematical principles, but because I regard as unreasonable in the extreme the 'conclusions' derived from an examination of real world behaviour using this mathematical methodology as a frame of reference.

1.5.2 My most serious disagreements are in the area of so-called stockmarket efficiency, where the practical results of well-documented actuarial models, such as Weaver & Hall (1967) and Clarkson (1981), are totally inconsistent with the findings of prominent financial economists such as Fama (1970) and Jensen (1968), and in the general approach to risk, which I regard as unsatisfactory in the extreme. As a consequence, the modern theory of finance has, I believe, followed the mistaken and unnecessarily fatalistic path of assuming that the real financial world is characterised by some inescapable risk-adjusted equilibrium position, rather than striving to find better ways of achieving personal and corporate objectives in the face of future uncertainties.

1.5.3 I must stress, however, that I make no personal criticism of those whom I perceive to be living in the other scientific world. I regard it as not unreasonable, in the light of their training and experience, for them to hold to the scientific views that they do.

1.6 *The Further Rise of Financial Economics*

Two developments, above all others, greatly enhanced the credibility of financial economics from the mid 1970s onwards. The first was the conceptual breakthrough of the Black-Scholes (1973) option pricing formula, which very quickly became established as a methodology of immense practical usefulness in the burgeoning new field of derivative securities. The second was the award of the 1990 Nobel Prize for Economics to Markowitz, Miller and Sharpe for their pioneering work in the theory of financial economics.

1.7 *The Formation of AFIR*

In the late 1980s, the actuarial profession responded to the increased importance of financial risk modelling by setting up AFIR (Actuarial Approach for Financial Risks) as a section of the International Actuarial Association. The new section was formally approved in 1989, and its first international colloquium was held in Paris in 1990. However, although five international colloquia have now been held under its auspices, no consensus approach to risk has emerged which might form the nucleus of an actuarial paradigm that could be put forward as an alternative to the financial economics methodology.

1.8 *A Scientific Discussion Meeting in Edinburgh*

In April 1992, a discussion meeting on recent developments in the theory and practice of finance was held at the Royal Society of Edinburgh. The panel of speakers included leading figures such as Professor Robert Merton of Harvard University, Professor Barr Rosenberg, who now runs an investment management company, Andrew Rudd, co-author of textbooks such as Rudd & Clasing (1982) and Jarrow & Rudd (1983), and Myron Scholes, of Black-Scholes fame. Three general themes that arose in the presentations were very much closer to my practical viewpoint than to standard finance theory. First, areas of inefficiency were more widespread than had been previously thought. Second, practical risk evaluation was carried out by downside simulations involving possible adverse scenarios, rather than through the use of a symmetric measure such as variance over the entire range of outcomes. Third, there was a realisation that linear models, while elegant in mathematical terms, had some highly unsatisfactory features, and certain types of non-linear model (such as ARCH and GARCH) were suggested as possible alternatives.

1.9 *A Scientific Discussion Meeting in London*

1.9.1 In November 1993, a public discussion meeting on the subject of mathematical models in finance was held at the Royal Society of London, with the keynote speaker being Professor Harry Markowitz. The most interesting feature to me was the description that Professor Markowitz gave of the portfolio selection system for Japanese equities that he had recently helped set up for a Japanese securities house. In particular, he adopted a downside approach to risk by using the semi-deviation below the broad market index, and he introduced upper limits on share weightings rather than allowing these to be determined by a standard portfolio selection algorithm.

1.9.2 Following references by several speakers to examples of stockmarket inefficiency and to doubts about the relevance of existing theory, I made the following remarks in a general discussion session:

> "What is the best way forward for the theory of finance and the role that mathematics can play? To apply the mathematical methods discussed we need the best possible understanding of the 'hidden variables' behind security prices. Professor Ziemba's comments on how errors in expected values in portfolio selection models are far more serious than errors in covariances confirm this. But how should we make these estimates?
>
> We have two main choices. We can — as suggested by Edgar Peters in his book *Chaos and order in the capital markets* — invoke chaos theory, and try to find ever more complex mathematical relationships between security prices and other factors. This approach is unlikely to be of much assistance. The alternative is to go back to the basics of understanding the human mind and find a simple model of human behaviour to replace the expected utility maxim that is a central feature of economists' definitions of rational behaviour."

1.9.3 The last presentation of the second of the two days discussed academic theory in the area of dynamic asset allocation. In response to what I perceived to be the unhelpful nature of academic theory in this important practical area, I commented as follows in the final discussion session:

> "Experience suggests that an academic analysis of risk is futile in most of the situations. We need to concentrate, instead, on investor psychology; what Keynes described as 'anticipating the anticipations of others'. As opposed to the second order differential equations we have been discussing, I regard the behaviour of 'good' investors as being at least fourth order in nature, but we simply do not yet have such high order mathematical models. My guess is that, in addition to the aggregate uncertainty introduced at the portfolio level by Professor

Markowitz more than 40 years ago we need to build in uncertainty at the microscopic level along the lines of the Heisenberg Uncertainty Principle in quantum theory if we are to explain how individual investors behave and thereafter build up a new theory of finance on sounder foundations than those provided by the linear methodologies of the current theory."

1.9.4 I then illustrated these somewhat abstract observations with a practical example of the type of real world behaviour that could not be explained by existing theory:

"There is a story that emphasizes the two crucial components of successful investment: obtaining important information before your competitors, and understanding investor psychology. When the Battle of Waterloo was fought in 1815, there were no Reuters or Topic screens to give a real-time account of the engagement. Rothschild, however, had his own private sources of intelligence, and when his messenger brought him the crucial news at about 11 a.m. he went into the gilts market and sold heavily. Others assumed that we had lost at Waterloo, and the market plunged in waves of panic selling. Rothschild bought back at the bottom of the panic and then went out to lunch. When the Government messenger arrived in the afternoon it was realized that we had indeed won the Battle of Waterloo, and prices soared spectacularly. The enormous profits that George Soros made out of currency speculation around the time of sterling's ignominious exit from the Exchange Rate Mechanism suggest that investor psychology has not changed much over the past two or three centuries."

1.10 *The Widening Gulf between Financial Theory and Financial Practice*

1.10.1 What I had perceived to be the failure of finance theory even to keep up with, let alone to lead, finance practice had also been apparent to other investment actuaries. In particular, one of the themes for papers suggested by the Scientific Committee of the Fourth International Colloquium of AFIR in Orlando in April 1994 was 'How to bridge the widening gap between finance theory and finance practice'. The concluding paragraph of Clarkson (1994), which I wrote under that theme, is as follows:

"The empirical evidence set out in this paper is strongly in support of the conclusion that three cornerstones of the current theory of finance — the expected utility maxim, the Sharpe Diagonal Model, and the use of variability of return as a proxy for risk — can have no place in any new theory if the gulf between financial theory and financial practice is to be bridged. The discovery of this new theory will require a revolution in thinking as far-reaching as that which occurred in Astronomy during the Copernican Revolution. The empirical work described in this paper suggests in very general terms where this revolution may begin — painstaking descriptive work involving not only the precise nature of new information but also the manner in which investors of differing degrees of skill and experience respond to new information."

1.10.2 I was, by then, fairly certain that the most important prerequisite to the construction of any better theory of finance would be a much more realistic framework for the behaviour of fallible human beings than that provided by classical utility theory as the rationale for 'rational behaviour'.

1.11 *Bühlmann and Individual Rationality*

1.11.1 In June 1995, Professor Hans Bühlmann gave an invited lecture in

London, in which one of his main themes was that the assumption of 'individual rationality' is the weakest link in modern finance theory. If the actuarial premium for a particular insurance is P_A, and incorporates not only the expected cost of claims and the expenses of management, but also a target rate of return on capital employed, then this will usually differ from P_M, the general level of premium in the market. The obvious insight, that business should only be written if P_M is at least equal to P_A, is contrary to the teachings of finance theory, where the assumption of 'individual rationality' is used to 'prove' that P_M is always the correct underlying premium.

1.11.2 I regard Professor Bühlmann's comments as an exceptionally reassuring verification of the view that I had already expounded, namely that the key to a better theory of finance will be a much more realistic description of human behaviour than that implied by the 'rational behaviour' paradigm.

1.12 *Beginning the Search*

1.12.1 The paper falls into four main parts, with the first part building some essential foundations for the new framework. Since I am in no doubt that the key to further progress is a more realistic model of human behaviour, I discuss, in Section 2, the workings of the human mind, and place particular emphasis on two key themes — Sir Roger Penrose's fairly recent refutation of the 'algorithmic thought' paradigm, and Adam Smith's much earlier demonstrations of the human intellect's quest for philosophic systems which 'soothe the imagination' in the face of risk and uncertainty. In the light of these considerations, I then show, in Section 3, that the 'rational behaviour' paradigm cannot be an appropriate guiding principle.

1.12.2 To begin filling the vacuum left by my rejection of all systems of economic science predicated on the 'rational expectations' approach, I turn, in Section 4, to Tolstoy's *War and Peace*, since there are, I believe, strong parallels between Tolstoy's motive for writing *War and Peace* and my objective in writing the present paper.

1.12.3 After a very brief Section 5 on the characteristics of actuarial science, I describe, in Section 6, various analogies with the medical and physical sciences that will guide the construction of my new framework.

1.13 *Ten Actuarial Papers*

The second part builds on the foundations of the first, using commentaries on my ten main actuarial papers to identify important aspects of real world behaviour that will have to be incorporated within the general framework of the new approach. In the cases of the three papers dealing with risk, namely Clarkson & Plymen (1988), Clarkson (1989) and Clarkson (1990), I have incorporated within the commentaries some new material of a theoretical nature.

1.14 *Ten Further Pointers*

The third part begins with a review of the simplifying assumptions that underly

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

three areas of economic science which involve human choice under conditions of uncertainty and risk — utility theory, the theory of games, and portfolio theory. After explaining my viewpoint on stockmarket efficiency in the light of my conclusions on the practical relevance of these three bodies of theory, I then develop six quite different miscellaneous themes to provide the remaining building blocks required to complete my new framework.

1.15 *Practical Applications*

Since my predominant theme is the need for a new and much better model for human behaviour in the face of uncertainty and risk, the fourth and final part of the paper begins with some general comments on what I describe, in the context of my new framework, as 'rational financial behaviour'. I then describe some practical methodologies that follow very naturally from this new paradigm in three important fields of actuarial endeavour — investment management, asset/liability management, and capital projects. The final application of my new conceptual framework is to summarise my personal views on the explanatory power and practical relevance of different elements of financial economics, in terms of a two-dimensional classification, analogous, in many ways, to the periodic table in chemistry.

1.16 *General Remarks*

1.16.1 Despite being a mathematician by training, I regard the basic ideas in economic science in general, and in financial economics in particular, as being of vastly greater importance than the development of elegant mathematical formulae and equations. I have, therefore, endeavoured to write this paper in plain English as far as possible. The occasional lack of rigour seems a small price to pay for making the basic issues much more easily understood than would have been the case had numerous passages of impenetrable mathematics been included.

1.16.2 There are many more quotations than is normally the case in papers presented to sessional meetings, for two reasons. First, it was only relatively late in my business career that I came across many items which, although in my opinion of considerable practical interest, were very seldom mentioned (if at all) in the 'academic' literature and had never (to my knowledge at least) appeared in a course of reading for the actuarial examinations. Making relevant extracts from these items more accessible may reduce the length of time it takes other actuaries to become clear, in their own minds, about the respective merits of differing conceptual approaches. Second, the literature of economic science is now so voluminous that earlier writers of considerable eminence may have suggested, possibly some time ago, points of view that give greatly enhanced coherence and academic respectability to ideas which we may have formulated in the context of our own personal experiences, and which we might mistakenly believe are completely new. In this regard, I make numerous references to some of the lesser known writings of Adam Smith, Keynes and Allais.

2. Penrose, Adam Smith and a Robinson Crusoe Paradox

2.1 *Introduction*

My starting point is a firmly held conviction that human behaviour in the financial world, in the face of risk and uncertainty, is rarely based on what, broadly speaking, could be described as the intelligent mental processing of known information. The general incidence of what might be called 'financial accidents' appears to me to be several orders of magnitude higher than what would be expected if everyone acted 'rationally', in the sense of some scientifically correct processing of known information in conjunction with mathematics and logic in the pursuit of clearly identified objectives.

2.2 *Penrose and the Human Mind*

2.2.1 In 1989, Professor (now Sir) Roger Penrose published *The Emperor's New Mind*, in which he challenged the conventional wisdom that the human mind acts in a slavishly algorithmic manner along the lines of a very powerful computer. Although this "most engaging and creative tour of modern physics, cosmology, mathematics and philosophy that has ever been written" (to quote the *New York Times* book review) is, in many places, highly technical, Penrose's conclusion is a very simple idea, namely that the purely computational viewpoint of the human mind is untenable. He also expresses his belief that "it is through science and mathematics that some profound advances in the understanding of mind must eventually come to light".

2.2.2 The 'rational behaviour' paradigm, on which most economic models of human behaviour are based, generally assumes that 'known information' can be processed 'intelligently' by the human mind at great speed and with unfailing accuracy, and so is a special case of the purely algorithmic paradigm of the human mind that has, in my view, been so convincingly refuted by Penrose.

2.2.3 The combination of Penrose's pioneering work and my own observations of the financial world leaves me in no doubt that the conventional 'rational behaviour' approach cannot be used to construct models that will have relevance in the real financial world.

2.3 *Adam Smith and the Human Mind*

2.3.1 To begin filling the vacuum left by my rejection of the conventional 'rational expectations' approach, I shall go right back to Adam Smith — not, however, to his *Wealth of Nations*, which was published in 1776, but to his *History of Astronomy*, which was probably written shortly before 1758, although not published until 1795, some five years after his death.

2.3.2 Since relatively few people have even heard of, let alone read, this philosophical essay, the full title of which is *The Principles which Lead and Direct Philosophical Enquiries; Illustrated by the History of Astronomy*, it may be useful to begin with a few extracts from Heilbroner's introduction to *The Essential Adam Smith* (1986):

> "The main body of the essay recounts in considerable detail the changing conceptions of the heavenly order, from Ptolemy's hypothesis of concentric spheres, through the struggles of Kepler and Copernicus to explain the vagaries of planetary motion, to the summative achievement of Newton. As such, it is an excellent review of the historical development of a science, reminiscent, in its emphasis on the succession of one temporarily successful account by another, of Thomas Kuhn's well-known modern treatment of science as a succession of vulnerable 'paradigms' or conceptual frameworks.
>
> This is not the reason, however, that I have included a reading from the astronomy essay. It is because the review of astronomical theory is prefaced by another remarkable demonstration of Smith's psychological penetration. Before he settles down to his historical account of astronomical theory, Smith asks a disconcerting question: why do men theorize in the first place, whether about astronomy or anything else? Smith's answer, like his discussion of the superiority of the number three over the number four in the architecture of argument, draws on his sense of deep-seated psychological drives and resting points. Men theorize, he says, because they are impelled to set at rest the 'tumult' they feel in the face of new observations. It is not an abstract drive for truth that impels the search for theory, but the concrete promptings of anxiety,"

and

> "Smith's view does not mean that scientific inquiry cannot lead toward truth. It means, rather, that criteria quite different from truth have an important, often overlooked, part to play in guiding the course of scientific (Smith would say 'philosophical') argument."

2.3.3 After describing our need to classify and categorise new information to lessen the 'tumult' in our mind, Smith explains the disconcerting mental processes that occur when some new item defies classification:

> "The memory cannot, from all its stores, cast up any image that nearly resembles this strange appearance. If by some of its qualities it seems to resemble, and to be connected with a species which we have before been acquainted with, it is by others separated and detached from that, and from all the other assortments of things we have hitherto been able to make. It stands alone and by itself in the imagination, and refuses to be grouped or confounded with any set of objects whatever. The imagination and memory exert themselves to no purpose, and in vain look around all their classes of ideas in order to find one under which it may be arranged. They fluctuate to no purpose from thought to thought, and we remain still uncertain and undetermined where to place it, or what to think of it."

2.3.4 Smith then introduces the dimension of time into this soothing mental trait of pattern recognition. When two distinct events have been observed to follow each other in the past, the mind runs ahead of the physical reality in terms of anticipating similar behaviour in the future. If subsequent behaviour is, indeed, consistent with the past, these occurrences are seen as 'normal', and virtually no mental energy is required to process them:

> "The ideas excited by so coherent a chain of things seem, as it were, to float through the mind of their own accord, without obliging it to exert itself, or to make any effort in order to pass from one of them to another."

2.3.5 Should subsequent behaviour be inconsistent with the past, however, a quite different sequence of mental activity occurs:

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

"The fancy is stopped and interrupted in that natural movement or career, according to which it was proceeding. Those two events seem to stand at a distance from each other; it endeavours to bring them together, but they refuse to unite; and it feels, or imagines it feels, something like a gap or interval between them. It naturally hesitates, and, as it were, pauses upon the brink of this interval; it endeavours to find out something which may fill up the gap, which, like a bridge, may so far at least unite those seemingly distant objects, as to render the passage of the thought between them smooth, and natural, and easy."

2.3.6 It is at this point that Smith introduces his 'invisible chain' metaphor for science or philosophy, to describe how the human intellect strives to fill the perceived gap in our comprehension of the real world:

"The supposition of a chain of intermediate, though invisible, events, which succeed each other in a train similar to that in which the imagination has been accustomed to move, and which link together those two disjointed appearances, is the only means by which the imagination can fill up this interval, is the only bridge which, if one may say so, can smooth its passage from the one object to the other."

The mental anxiety caused by apparently perplexing behaviour is thereby removed:

"Upon the clear discovery of a connecting chain of intermediate events, it vanishes altogether. What obstructed the movement of the imagination is then removed. Who wonders at the machinery of the opera house who has once been admitted behind the scenes?"

2.3.7 Smith, however, was well aware that in the natural world the discovery of a clear 'connecting chain' was the exception rather than the rule, and he cited astronomy as one of the few 'success stories'. In other areas, where 'some coherence to the appearance of nature' had been achieved, he suggested that the discovery of even a partially successful connecting link is a step in the right direction.

2.3.8 Just before proceeding to his chronicle of how astronomy developed as a science, Smith generalises about philosophy (or, as we would call it now, the philosophy of science) as an intellectual discipline which possesses both theory and history in its own right, and is therefore worthy of our study. At any particular time, Smith observes that the best available scientific explanation may fall well short of the underlying 'truth'. Furthermore, a scientific theory that might be classified in Penrose (1989) as 'superb' (his other categories being 'useful' and 'tentative') could give the correct results, but for the wrong reasons. These very perceptive observations are at odds with the commonly held belief that 'science' is both infallible and unchallengeable.

2.3.9 Heilbroner explains how Smith's main conclusion is that the workings of the human mind have evolved in such a way as to offer limited mastery of behavioural patterns which are 'intelligent' in the face of risk and uncertainty:

"As Smith shows, much of the work of philosophers in the historical evolution of astronomy has been an effort to lessen the anxiety of Wonder by placing observations within a reassuring 'system' of ideas. Thus the real aim of science goes beyond the attempt to penetrate the

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

recesses of nature. Science seeks to help us to live in a universe where Surprise and Wonder threaten to undo our necessary equilibrium, our essential peace of mind."

2.4 *The Human Mind and the Constraint of Time*

2.4.1 Two crucial differences between the algorithmic explanation of human thought, and the 'Adam Smith' explanation, as outlined above, both involve the dimension of time. Firstly, the human mind is many orders of magnitude faster at pattern recognition than at algorithmic operations of a numerical or logical nature. Secondly, the human mind is relatively slow — in comparison to a normal human lifespan — at assimilating the training and experience required to operate 'intelligently' in the face of risk and uncertainty.

2.4.2 An obvious corollary is that any problem-solving approach which uses algorithmic thought in the application of some 'philosophic system' will take a finite, and possibly considerable, amount of time that will vary with the abilities of the individual. This dependence on abilities is completely absent from all the conventional formulations of 'rational behaviour'. A second corollary is that it may take a particular individual so long to find the solution that the opportunity being analysed no longer exists, having been spotted and acted upon more promptly by others. The potential for achievement is thus very strongly dependent on the relative abilities of the individual, rather than being an inherent and invariant characteristic of the external world.

2.4.3 My Robinson Crusoe paradox highlights the acute difficulties that can be caused by the constraint of time. First, however, I shall illustrate some of the essential differences between algorithmic thought and pattern recognition thought by means of a very elementary computational problem, namely finding the highest common factor of two fairly small numbers.

2.5 *Highest Common Factors*

2.5.1 One of the Ancient Greeks' main contributions to philosophy and science was the development of a rigorous framework for the arithmetic of integers, and an excellent example of a practical technique is Euclid's Algorithm for finding the highest common factor of two numbers. This ingenious algorithm involves the successive subtraction of multiples of the smaller of two numbers from the larger number until the remainder is zero. For example, for 725 and 2,436 the steps are as follows:

$$
\begin{aligned}
2{,}436 &= 3 \times 725 + 261 \\
725 &= 2 \times 261 + 203 \\
261 &= 1 \times 203 + 58 \\
203 &= 3 \times 58 + 29 \\
58 &= 2 \times 29 + 0
\end{aligned}
$$

so that the highest common factor is 29.

2.5.2 You are now asked two questions involving mental arithmetic:

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

(1) what is the highest common factor of 22 and 77? and
(2) what is the highest common factor of 21 and 91?

2.5.3 In the first case, the answer of 11 is seen at a glance, since the necessary steps "float through the mind of their own accord, without obliging it to exert itself", to use Smith's words. The pattern of two identical digits is immediately recognised as a multiple of 11, and the other factors, namely 2 and 7, have no factor higher than 1 in common.

2.5.4 In the second case, obtaining the answer of 7 takes much longer, and requires some mental exertion, along the following lines:

(a) 21 is immediately seen to be 3 x 7, but a factorisation of 91 does not come to mind unaided.
(b) We must test whether 3 or 7 divides 91.
(c) 91 is not a multiple of 3, since clearly 90 is a multiple of 3.
(d) 7 into 9 gives remainder 2; carrying 2 gives 21, which is a multiple of 7.
(e) 91 is a multiple of 7.
(f) The highest common factor is 7.

2.5.5 In the second case, the number 91 'stands alone and by itself in the imagination', since the multiplication tables which we had to learn by rote at primary school stopped at the '12-times table'. Step (b), the insight required before effective further progress is possible, may not be obvious to someone with no formal training in mathematics, and steps (c) and (d) require training and competence in arithmetic. We therefore expect that the time taken in this second case will vary significantly from individual to individual, depending both on what might be called innate intelligence and on the level of competence achieved in arithmetic through practical experience.

2.5.6 It would, of course, be possible to use Euclid's Algorithm in the second case. However, although the essential working can be set down very elegantly on paper in only two lines:

$$91 = 4 \times 21 + 7$$
$$21 = 3 \times 7 + 0$$

both the time and the mental energy required, first of all to recall the rules of the algorithm, and then to apply them to this case, are far greater than for the *ad hoc* problem-solving approach, which involves a combination of pattern recognition thought and basic arithmetic.

2.6 *Robinson Crusoe's Choice under Uncertainty*

2.6.1 Robinson Crusoe has built a large raft and sailed it to a nearby island, which he expected would be uninhabited. On landing at a sandy beach, he is suddenly confronted by three natives, Allan, Barry and Cameron. Each is either a Truthteller (who always tells the truth), a Falsifier (who always lies), or an Alternator (who alternately tells the truth and lies). One of them is Prime Minister of the island.

2.6.2 Allan says:
(1) The Prime Minister is of a different type from each of the other two of us.
(2) Barry is not the Prime Minister.

Barry says:
(1) The Prime Minister is a Falsifier.
(2) Allan is not the Prime Minister.

Cameron says:
(1) Exactly two of us are of the same type.
(2) I am not the Prime Minister.

2.6.3 If Robinson Crusoe can work out within 15 minutes who is Prime Minister, he can take back to his own island, on this and subsequent visits, as much of the plentiful food and other natural resources that he could ever want. If he accepts this challenge, but fails to arrive at the correct answer within 15 minutes, he will be killed instantly. If he chooses not to accept the challenge, he will be allowed to leave the island unharmed. He has one minute in which to decide whether or not to accept the challenge.

2.6.4 What should Robinson Crusoe do?

2.7 *Levels of Understanding*

2.7.1 Let us begin by classifying individuals into four broad groups, the first of which represents the not insignificant number of people for whom this problem is well beyond their intellectual powers, no matter how much time is allowed. They could only arrive at the correct answer by pure guess-work, which gives the dangerously high probability of 2/3 of obtaining the wrong answer.

2.7.2 The next group comprises those who have a very limited understanding of the characteristics of such a problem. They might mistakenly believe that there are only 3 cases to check out (i.e. either Allan, Barry or Cameron is the Prime Minister and the other two are not) and that 15 minutes would provide ample time to test which of these 3 mutually exclusive possibilities is consistent with the 6 statements.

2.7.3 The third group represents those who have a limited familiarity with logic problems, but no great practical experience in solving them. They will realise that the solution can probably only be found by determining what types the 3 natives are. They may mistakenly think that there are only 3 x 3 x 3, i.e. 27, cases to check out, namely each native being either a Truthteller, a Falsifier, or an Alternator. The allotted time of 15 minutes might then be thought sufficient to work through these 27 cases in turn to find which provides the solution.

2.7.4 The fourth and final group comprises those who are well aware of the general characteristics of such logic puzzles, either through extensive practical experience or through a formal course of tuition in mathematical logic, possibly as part of an honours course in mathematics at a university. To this group it will

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

be obvious that the 'Alternator' classification has to be sub-divided into two distinct sub-groups, depending on whether it is the first or the second of the two statements that is true. The number of possible cases is thus 4 x 4 x 4 i.e. 64, which they will realise would probably take longer than 15 minutes to check through in slavish algorithm fashion using a 'truth table' to find which (not necessarily unique) assignation of types is consistent with the six statements, and, therefore, corresponds to the solution.

2.7.5 When I first came across this logic puzzle on its own (i.e. without my embellishment of the 15 minute 'guillotine'), I tackled it in a very leisurely fashion, using what I would call 'case elimination by intelligent guess-work'. Since statement C(1) (in an obvious notation) seemed, if true, to narrow down the possibilities in a readily understandable way to a greater extent that either A(1) or B(1) being true, I first of all assumed that C was a Truthteller, and was able to show that this led to a contradiction, and hence could not be the case. Similarly, I showed that the assumption that B was a Truthteller led to a contradiction. Then I showed that if none of A, B or C was a Truthteller, a contradiction again resulted, from which I could deduce that A was a Truthteller. It was then quite straightforward (not unlike an easy 'set-piece' end-game in chess) to deduce that B was a Falsifier and C was an Alternator, so that, in particular, B(2) was false, and Allan was, accordingly, the Prime Minister.

2.7.6 I should have been able to arrive at this solution in, perhaps, 10 minutes, but I made a mistake in overlooking one possible case about half-way through. As a result, I wasted around 5 minutes retracing my steps, and correcting this error, once I arrived at the clearly erroneous conclusion that no solution existed. Just as I was finishing the problem after about 15 minutes, I realised from the symmetry of statements A(2), B(2) and C(2) that precisely two of these had to be true, with the other false, and that, for each of these 3 primary cases, the maximum of 8 secondary cases (i.e. A(1), B(1) and C(2) each either true or false) could be resolved very quickly, so that — given this insight — the solution should be reached in around 5 minutes at the most.

2.8 *Uncertainty of the First Kind*

2.8.1 It is now obvious that what might be called the 'par for the course' solution time will vary with the approach adopted, from around 5 minutes, given the insight described above, to perhaps 20 or 25 minutes for an algorithmic truth table approach, with the *ad hoc* approach that I initially used coming somewhere in the middle of the range, at around 10 or 15 minutes. By analogy with golf, I see my 'par' solution time as a fairly demanding target for an expert. Better performances are possible, given either exceptional skill or sheer good luck, but in real life human fallibility and the tendency for errors to compound in severity lead to poorer performances being much more likely.

2.8.2 I shall describe variations in my 'par' solution time as 'uncertainty of the first kind'. Such variations relate to the most fundamental underlying source of lack of certainty in any estimation of the likely solution time.

2.9 *Other Uncertainties*

The most obvious other source of lack of certainty in any estimation of the likely solution time is human error. The scope for errors of this nature will vary considerably from one individual to another.

2.10 *A Common Sense Strategy*

2.10.1 I can now give my somewhat lengthy answer as to what Robinson Crusoe should do. If he falls into one of my first three categories (which I shall call 'totally inexperienced', 'beginner' and 'intermediate' respectively) he should reject the challenge, leave the island immediately, and live to fight another day.

2.10.2 If he is in my fourth category (which comprises those who could be described as 'advanced' or 'expert') he should, first of all, realise fairly quickly that a 'truth table' approach has virtually no chance of success, and that it will probably take of the order of 10, 15 or 20 minutes to solve the problem in an *ad hoc* manner, which would result in a dangerously high probability of not being able to arrive at the correct solution in time. The remainder of the one minute thinking time should then be spent looking for an ingenious approach which is likely to yield the solution in around 5 minutes or so, which would leave an acceptably low chance of failure.

2.10.3 If he believes that he sees a promising way forward and, accordingly, accepts the challenge, two cases arise. If he reaches the solution within 15 minutes, all is well. If, shortly before the 15-minute 'guillotine' falls he has not reached the solution, he must spend the little remaining time in a damage limitation exercise to assess which of the three natives has the highest likelihood of being Prime Minister. If, for example, he is fairly sure that Cameron is not the Prime Minister and has also a 'gut feel' that it is more likely to be Allan than Barry, he should state 'Allan', and tactfully omit mentioning the incompleteness of his analysis. Here, as with many real life situations, reward or punishment depends purely on whether, with hindsight, the decision proves to have been 'correct' or not.

2.11 *Elements of Paradox*

The paradox is that this Robinson Crusoe situation, which will be recognised as being not dissimilar, in principle, to some of the high-risk situations that arise in the real financial world, highlights various self-evident features that are inconsistent with the teachings of models based on the 'rational behaviour' paradigm. The main such elements of paradox are:

(1) Individuals with insufficient prior experience or training will tend to understate, perhaps to a very dangerous degree, the true level of risk.

(2) Determination of the appropriate strategy involves the application of pattern recognition thought to prior experience, rather than being dependent solely on the application of algorithmic thought to the current situation.

(3) The course of action (i.e. accept or reject the challenge) to which the strategy will lead is indeterminate, in that it is impossible to predict in advance

whether or not an approach involving sufficiently low risk will be spotted in time.

2.12 *The Utility Theory Approach*

2.12.1 Let us now suspend, for the moment, the constraint of time, and allow Robinson Crusoe to call up two professional advisers on his satellite-linked mobile telephone — a financial economist and an actuary who specialises in stochastic modelling. Both are asked to send back to his satellite-linked fax machine a practical course of action to allow him to decide whether or not to accept the challenge. The fax from the financial economist is as follows:

> "The obvious first step is to produce, from your previous experience of similar logic puzzles, the probability density function for the time that it will take you to solve this particular puzzle. While much of the theory built up by financial economists makes use of the familiar bell-shaped normal or Gaussian distribution, it is likely, in this case, that the distribution will be skewed to the upside, and you should therefore use a lognormal distribution, which is a well-known standard distribution. In assessing the risk of accepting the challenge, two of the general principles of mean-variance analysis are relevant. First, in this case a lower mean or average value is 'good' in that — other things being equal — it reduces the probability of your time exceeding 15 minutes. Second, for a given value of the mean, the higher the variance (or, equivalently, the standard deviation) the higher will be the probability of your time exceeding 15 minutes. Uncertainty, as measured by variance, is therefore 'bad'. You should now consider very carefully the personal utility function that you employ, in risk-reward situations like this, to combine the mean and the variance into one numerical measure. When you have obtained a numerical value for the utility of accepting the challenge, you should compare it with your utility value for declining the challenge and, thereby, missing out on the opportunity of a massive enhancement to your present quality of life, albeit at a possibly significant risk to life. If the former utility value is the larger you should accept the challenge; otherwise you should reject it."

2.12.2 I see four flaws in this strategy, three serious and the fourth irremediable. First, it does not recognise in any way what I call 'uncertainty of the first kind', namely that the crucial judgement relates to the 'true' underlying difficulty, and hence the 'par for the course' solution time. Second, someone who is not highly experienced in solving similar logic puzzles will tend to underestimate very considerably the likely time to solve the puzzle, and, thereby, systematically understate very significantly the inherent risk. Third, the degree of skewness — a crucial variable factor for the lognormal class of distribution — has not been discussed, even although the higher the skewness to the right the higher, for the same mean and variance, will be the risk. The fourth, and totally irremediable, flaw is that most individuals are totally incapable of specifying their 'personal utility function', even in low-risk situations.

2.13 *A Stochastic Model*

2.13.1 The fax received from the actuary specialising in stochastic models reads as follows:

> "By good fortune I have had access to the results of all your previous attempts at solving logic puzzles of precisely this level of difficulty. After studying the statistical properties of the crude

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

data in accordance with generally accepted principles, I have concluded that the most appropriate stochastic model is as follows:

Let x be a random variable which takes each of the values 0, 1, 2, 3, ... 8 and 9 with equal probability, and let N(0,1) be a unit normal random variate. Then the time $t(x)$ in minutes to solve the puzzle is given by:

x	$t(x)$
0,1	2 (N(0,1) + 2.5)
2,3	4 (N(0,1) + 2.5)
4,5,6,7	6 (N(0,1) + 2.5)
8,9	8 (N(0,1) + 2.5)

To allow you to run this model without a calculator or computer of any type, I have appended copies of the necessary pages from *Actuarial Tables for Examination Purposes*, namely the page of random numbers and the page containing the data from which you can extract the approximate percentile points of the unit normal distribution. By using these to run the model, you should find that the probability of your taking more than 15 minutes to solve the puzzle will converge to a limit after, say, 100 or 200 trials. You should then accept or reject the challenge, depending on whether or not this value is respectively below or above the probability of imminent death that you are prepared to accept in the light of the very significant potential rewards available."

2.13.2 Fortunately Robinson Crusoe is a competent mathematician and can, therefore, implement the actuary's instructions. First of all he uses the values of the distribution function F(y) of the unit normal distribution to produce the following table of 'percentile ordinates':

Percentile	F(y)	Ordinate
1	0.005	−2.582
2	0.015	−2.173
3	0.025	−1.96
—	—	—
50	0.494	−0.01
51	0.505	0.01
52	0.515	0.04
—	—	—
98	0.975	1.96
99	0.985	2.17
100	0.995	2.58

2.13.3 He then runs the stochastic model by taking triples of random digits from the table of random numbers. For example, the first triple of 885 gives a value of 8 for x, a percentile value of 85 which corresponds to an ordinate of 1.02, and hence a time $t(x)$ = 8 (1.02 + 2.5) i.e. 28.2 minutes. Proceeding thus, Robinson Crusoe builds up a schedule of which the first 5 lines are as follows:

Trial	Random digits	$f(x)$	$N(y)$	$t(x) = f(x) \times [N(y) + 2.5]$
1	8; 85	8	1.02	28.2
2	6; 53	6	0.06	15.4
3	2; 75	4	0.66	12.6
4	9; 33	8	−0.45	16.4
5	3; 57	4	0.16	10.6

2.13.4 After running the model for 100 trials, Robinson Crusoe draws up the following summary table:

	Frequency of $t(x)$ above 15		Probability of failure	
Total trials	Last 10 trials	Cumulative	Last 10 trials	Cumulative
10	6	6	0.60	0.60
20	4	10	0.40	0.50
30	2	12	0.20	0.40
40	2	14	0.20	0.35
50	5	19	0.50	0.38
60	6	25	0.60	0.42
70	6	31	0.60	0.44
80	5	36	0.50	0.45
90	4	40	0.40	0.44
100	2	42	0.20	0.42

Robinson Crusoe is now fairly confident that the probability of failure will tend to a limit very close to 0.42, in which case it is obvious that the challenge should be rejected out of hand.

2.14 *Making Sense of the Stochastic Model*

2.14.1 With a probability of failure of around 0.4, the course of action indicated by the stochastic model is an unqualified 'reject the challenge'. Robinson Crusoe cannot comprehend this conclusion in the light of his 'common sense' approach of accepting the challenge, provided that he is sufficiently confident that he can solve the puzzle well within the 15-minutes time limit. However, in trying to make sense of the model, he is seriously hindered by its lack of transparency. It is by no means obvious, from the essentially statistical description of the model, that it is even broadly appropriate as a framework for the very high-risk situation that he faces.

2.14.2 A more specific problem relates to the fact that the model has been fitted to the outcomes of all previous attempts at logic puzzles of the same degree of difficulty. Given the likely improvement in his success rate over time as his practical experience increases, he regards it as obvious that his more recent experience will be of far more relevance than data relating, in part, to his tentative early attempts at such puzzles.

2.14.3 The logical place to start is with the problem of the lack of transparency. The essentially statistical description, in Section 2.13, can be expressed in the following equivalent terms:

Probability	Solution time in minutes
0.2	5 + 2N(0,1)
0.2	10 + 4N(0,1)
0.4	15 + 6N(0,1)
0.2	20 + 8N(0,1)

where, as before, N(0,1) is a normal variate with an expected value of zero and a standard deviation of 1.

2.14.4 This, in turn, can be translated into a statement closer to plain English:

> "There are four possible scenarios, corresponding to the average solution time that corresponds to the effectiveness of the manner in which you choose to tackle the problem. There is a 20% chance that this average time is 5 minutes, a 20% chance that it is 10 minutes, a 40% chance that it is 15 minutes, and a 20% chance that it is 20 minutes. Whatever the expected time, the effect of essentially unpredictable factors will result in a scatter of actual times around the expected time. In accordance with conventional statistical methodology, this random error — which is assumed to be symmetric about the mean — is modelled using a normal distribution. Statistical analysis of the data suggests that a standard deviation of 40% of the expected time is appropriate."

2.14.5 Robinson Crusoe can now understand the rationale of the model in 'general reasoning' terms:

> "While, in theory, the average solution time, which depends on the effectiveness of the particular approach adopted, will vary continuously, a reasonable pragmatic approach is to consider — with appropriate probabilities — four specific scenarios, each of which can then be modelled in terms of mean and variance by a normal distribution."

2.15 *The Conflict with Common Sense*

The essential conflict is obvious; a strategy derived from common sense principles appears to be inconsistent with the results obtained from generally accepted applications of stochastic models.

2.16 *Conclusions*

Neither the utility theory approach of financial economics nor the stochastic modelling approach of actuarial science appears capable of giving a practical way forward in the resolution of Robinson Crusoe's high-risk choice under uncertainty.

3. Rational Behaviour

3.1 *Introduction*

In this section I investigate the general characteristics of three different interpretations of what economists call 'rational behaviour' — the 'Robinson Crusoe' model, the 'theory of games' approach developed by von Neumann & Morgenstern (1944), and the 'rational man' approach developed in Markowitz (1959), and then comment on their appropriateness or otherwise.

3.2 *Robinson Crusoe*

3.2.1 The Robinson Crusoe approach is described by von Neumann & Morgenstern as 'a pure maximisation problem' which presents only practical, not theoretical, difficulties, and has the following characteristics:

> "Apart from those variables which his will controls, Crusoe is given a number of data which are 'dead'; they are the unalterable physical background of the situation. (Even when they are apparently variable, cf. footnote 2 on p. 10, they are really governed by fixed statistical laws.)

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

Not a single datum with which he has to deal reflects another person's will or intention of an economic kind — based on motives of the same nature as his own."

3.2.2 The footnote referred to is:

"Sometimes uncontrollable factors also intervene, e.g. the weather in agriculture. These however are purely statistical phenomena. Consequently they can be eliminated by the known procedures of the calculus of probabilities: i.e., by determining the probabilities of the various alternatives and by introduction of the notion of 'mathematical expectation'."

3.2.3 My Robinson Crusoe paradox, in the preceding section, suggests that any approach which assumes 'known probabilities' and ignores the very limited mental powers of human beings in the area of algorithmic computation, is unlikely to be of much relevance.

3.3 *Allowing for the Others*

Von Neumann & Morgenstern reject economic models of the Robinson Crusoe type as being unable to take into account the likely actions of 'the others', and appreciate that it is inappropriate always to impute rationality to 'the others':

"In whatever way we formulate the guiding principles and the objective justification of 'rational behaviour', provisos will have to be made for every possible conduct of 'the others'. Only in this way can a satisfactory and exhaustive theory be developed. But if the superiority of 'rational behaviour' over any other kind is to be established, then its description must include rules of conduct for all conceivable situations — including those where 'the others' behaved irrationally, in the sense of the standards which the theory will set for them."

3.4 *Simplifying Assumptions*

However, this ambitious objective is very seriously weakened by the crucial assumptions that they are forced to make:

"Subsequent discussions will show that we cannot avoid the assumption that all subjects of the economy under consideration are completely informed about the physical characteristics of the situation in which they operate and are able to perform all statistical, mathematical, etc., operations which this knowledge makes possible. The nature and importance of this assumption has been given extensive attention in the literature and the subject is probably very far from being exhausted. We propose not to enter upon it. The question is too vast and too difficult and we believe that it is best to 'divide difficulties'. I.e. we wish to avoid this complication which, while interesting in its own right, should be considered separately from our present problem."

3.5 *The Common Individuals*

Again, the very narrow theory which results from these highly unrealistic assumptions is unlikely to be of much relevance. Indeed, von Neumann & Morgenstern themselves observe that:

"Evidently the common individual, whose behaviour one wants to describe, does not measure his utilities exactly but rather conducts his economic activities in a sphere of considerable haziness."

3.6 *Portfolio Theory*

In Markowitz (1959) the section on the theory of rational behaviour begins as follows:

> "A portfolio analysis is characterised by
> (1) the information concerning securities upon which it is based;
> (2) the criteria for better and worse portfolios which set the objectives of the analysis; and
> (3) the computing procedures by which portfolios meeting the criteria in (2) are derived from the inputs in (1).
>
> The results of a portfolio analysis are no more than the logical consequences of its information concerning securities. Although they may not be apparent to unaided reason, the results, nevertheless, are but restatements of these inputs."

3.7 *Rational Man*

Unlike von Neumann & Morgenstern, Markowitz does *not* make the assumption of 'perfect information':

> "The theory of rational behaviour is usually presented as a study of the principles upon which a rational man would act. This Rational Man is unlike you or me in that he makes no errors in arithmetic or logic in attempting to achieve his clearly defined objectives. He is like you and me, on the other hand, in that he is neither omnipotent nor omniscient. He must make decisions, such as the selection of a portfolio, in the face of uncertainty. Since his information is limited, he may take less than perfect actions. Since his powers are limited, his achievements may fall short of the best conceivable. Every action, however, is perfectly thought out; every risk is perfectly calculated."

3.8 *A New Viewpoint*

Again, unlike von Neumann & Morgenstern, Markowitz does *not* take into account the likely 'irrationality' of some investors, and accordingly puts forward his approach only as 'a new viewpoint' to complement common sense as a basis for making better judgements:

> "The Rational Man, like the unicorn, does not exist. An attempt to see general principles by which he would act, however, can be suggestive for our own actions.
>
> The theory of rational behaviour is not a substitute for human judgment. There is no integrated theory by which we could dispense with human beings if we had a sufficiently large and fast computer. The study of rational behaviour has produced only general principles to be kept in mind as guides. Even the significance of some of these principles is subject to controversy. The value of the study of rational behaviour is that it supplies us with a new viewpoint on problems of criteria — a viewpoint to be added to common sense to serve as a basis of good judgment."

3.9 *The Expected Utility Maxim*

Markowitz then explains how the expected utility maxim, which is central to his approach, can be justified by the axiomatic approach to utility theory as a framework for human choice under uncertainty, even although some commentators are strongly critical of such an approach:

> "The expected utility maxim appears reasonable offhand. But so did the expected return maxim when first used. The latter maxim, we saw, is quite unacceptable. Perhaps there is some equally strong reason for decisively rejecting the expected utility maxim as well.

In recent years a justification has been presented that goes beyond the apparent plausibility of the expected utility maxim. The new axiomatic approach begins with basic principles which seem beyond denial, then demonstrates that the expected utility maxim follows from these principles. The axiomatic approach has revived interest and gained a large number of adherents for the two hundred year old expected utility maxim.

Some recent commentators, on the other hand, have argued that the expected utility maxim is not the essence of rational behaviour. They show instances in which human action differs from that dictated by the maxim. More pertinent, they show instances where reasonable action and the expected utility rule apparently contradict. These opponents of the maxim claim that, while the axioms upon which the modern argument is based have immediate appeal, they conceal objectionable assumptions."

This axiomatic approach essentially says that the human mind acts in a slavishly and unfailingly accurate algorithmic manner, which is inconsistent both with Penrose's conclusions and with the alternative 'Adam Smith' model of 'pattern recognition thought' described in the preceding section.

3.10 *Contradictions*

Markowitz returns later to the contradictions raised by the application of the expected utility maxim to specific situations:

"The writer feels that the most interesting and relevant arguments against the expected utility maxim involve specific cases in which human subjects, after careful deliberation, choose alternatives inconsistent with the maxim. The situations are reasonably simple, the human choice fairly definite, the contradiction between choice and maxim apparently inescapable. Either we must conclude that the expected utility maxim is not the criterion of rational behaviour, or else we must conclude that the human being has a natural propensity toward irrationality, even in his most reflective moments."

3.11 *Irrational Behaviour*

Markowitz describes how he concludes that the apparently anomalous behaviour, cited by Allais, is, in fact, 'irrational behaviour':

"Three examples are given below in which it is possible for an individual to prefer an alternative inconsistent with the expected utility maxim. In each case the 'wrong' alternative has a plausible appearance and is chosen by many persons questioned. The first example is one of the author's. The other two are due to Allais. The conclusion which Allais drew from his examples is that, since reasonable men choosing among simple alternatives contradict the expected utility maxim, this rule must be a poor one. The conclusion which this writer drew was that the individuals choosing the 'wrong' alternative acted irrationally."

3.12 *Systematic Irrationality*

I shall now cite various items of evidence to show that what Miller & Modigliani (1961) called 'systematic irrationality' is such a prevalent feature of human behaviour that neither the von Neumann & Morgenstern formulation, nor any formulation based purely on some other interpretation of 'rational behaviour', can explain actual human behaviour in the real financial world.

3.13 *Adam Smith and Impediments in the Minds of Men*

3.13.1 The writings of Adam Smith, described in the preceding section, relate

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

to the 'early Adam Smith' and, in particular, to the philosophic systems he studied during his period of self-education at Oxford University. The 'mature Adam Smith' writes in his *Wealth of Nations* about the impediments in the minds of men that make 'rational' equilibrium positions impossible in the context of wage rates. In particular, based on his observations of human behaviour over many years in many countries, he observes:

> "The chance of gain is by every man more or less over-valued, and the chance of loss is by most men under-valued, and by scarce any man, who is in tolerable health and spirits, valued more than it is worth."

3.13.2 As supporting evidence, he cites the failure to take out fire insurance on houses or insurance on ships, even though, in the latter case, 'self-insurance' may be an intelligent strategy:

> "When a great company, or even a great merchant, has twenty or thirty ships at sea, they may, as it were, insure one another. The premium saved upon them all may more than compensate such losses as they are likely to meet with in the common course of chances. The neglect of insurance upon shipping, however, in the same manner as upon houses, is, in most cases, the effect of no such nice calculation, but of mere thoughtless rashness and presumptuous contempt of the risk."

3.13.3 One of the most important economic decisions in life is the choice of career, but, according to Smith, behaviour in this area often falls far short of 'optimal' or 'rational':

> "The contempt of risk and the presumptuous hope of success are in no period of life more active than at the age at which young people choose their professions."

If such important decisions are not addressed intelligently, there seems little hope for 'rational behaviour' to be a useful model in other areas of economic choice under uncertainty.

3.14 *Keynes and Human Behaviour*

3.14.1 In the highly perceptive Chapter 12 of his *General Theory*, Keynes suggests that it is virtually impossible to make sensible projections of the future:

> "Our knowledge of the factors which will govern the yield of an investment some years hence is usually very slight and often negligible. If we speak frankly, we have to admit that our basis of knowledge for estimating the yield ten years hence of a railway, a copper mine, a textile factory, the goodwill of a patent medicine, an Atlantic liner, a building in the City of London amounts to little and sometimes to nothing; or even five years hence. In fact, those who seriously attempt to make any such estimate are often so much in the minority that their behaviour does not govern the market."

3.14.2 Even the idea that prices might be 'correct on average' is emphatically dismissed by Keynes:

> "It might have been supposed that competition between expert professionals, possessing judgment and knowledge beyond that of the average private investor, would correct the

vagaries of the ignorant individual left to himself. It happens, however, that the energies and skill of the professional investor and speculator are mainly occupied otherwise. For most of these persons are, in fact, largely concerned, not with making superior long-term forecasts of the probable yield of an investment over its whole life, but with foreseeing changes in the conventional basis of valuation a short time ahead of the general public. They are concerned, not with what an investment is really worth to a man who buys it 'for keeps', but with what the market will value it at, under the influence of mass psychology, three months or a year hence. Moreover, this behaviour is not the outcome of a wrongheaded propensity. It is an inevitable result of an investment market organised along the lines described. For it is not sensible to pay 25 for an investment of which you believe the prospective yield to justify a value of 30, if you also believe that the market will value it at 20 three months hence."

3.14.3 Keynes describes the importance of 'confidence' and 'animal spirits' as factors which invalidate any quantitative basis for rational behaviour:

"The *state of confidence*, as they term it, is a matter to which practical men always pay the closest and most anxious attention. But economists have not analysed it carefully and have been content, as a rule, to discuss it in general terms;"

"Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as a result of animal spirits — of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities."

3.14.4 Keynes, however, acknowledges that we endeavour to act 'intelligently' in the absence of any definitive quantitative or probabilistic framework:

"We are merely reminding ourselves that human decisions affecting the future, whether personal or political or economic, cannot depend on strict mathematical expectation, since the basis for making such calculations does not exist; and that it is our innate urge to activity which makes the wheels go round, our rational selves choosing between the alternatives as best we are able, calculating where we can, but often falling back for our motive on whim or sentiment or chance."

3.15 *Pepper and the Flow of Funds*

Pepper (1994) is mainly devoted to describing various (often lagged) linkages between stockmarket prices and specific aspects of economic policy which have an effect on 'the flow of funds'. However, he also highlights instabilities that are totally inconsistent with any 'rational behaviour' explanation of capital market behaviour:

"Prices peak when economic news is bad; they respond only to good news when they are rising, or to bad news when they are weak; they overshoot, and then correct violently."

3.16 *Conclusions*

My conclusion is that 'rational behaviour', however defined, can never be the cornerstone for any theory which purports to explain human behaviour under uncertainty. We are then driven to address the much more perplexing question of how to encapsulate the characteristics of 'irrational behaviour'. As a starting point in this difficult quest, I discuss, in the next section, Tolstoy's observations on the

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

nature of the psychological and other forces that shape events and on the inability of the human intellect to understand these driving forces.

4. Tolstoy and Human Behaviour

4.1 *Introduction*

While many people regard Tolstoy's *War and Peace* as the greatest novel ever written, he himself saw his great work quite differently, and wrote as follows in the specialist periodical *Russian Archive* in 1868:

> "What is *War and Peace*? It is not a novel, even less is it a poem, and still less an historical chronicle. *War and Peace* is what the author wished and was able to express in the form in which it is expressed.
>
> For me the most important consideration relates to the small significance that, in my conception, should be ascribed to so-called great men in historical events. Studying so tragic an epoch, so rich in the importance of its events, so near to our own time, and regarding which so many varied traditions survive, I arrived at the evident fact that the causes of historical events when they take place cannot be grasped by our intelligence."

4.2 *Conceptual Errors*

Apart from the famous Epilogue which is devoted to a discussion of the essential conflict between freewill and necessity, the most important philosophic, rather than narrative, section of *War and Peace* is Chapter 1 and Chapter 2 of Part XI. The first paragraph of Chapter 1 sets out Tolstoy's main conclusion as to why the human mind makes conceptual errors:

> "For the human mind the absolute continuity of motion is inconceivable. The laws of motion of any kind only become comprehensible to man when he examines units of this motion, arbitrarily selected. But at the same time it is from this arbitrary division of continuous motion into discontinuous units that a great number of human errors proceeds."

4.3 *Achilles and the Tortoise*

4.3.1 Tolstoy then describes the paradox of Achilles and the tortoise, which could not be solved by the mathematics available to the Ancient Greeks:

> "We all know the so-called sophism of the ancients, proving that Achilles would never overtake the tortoise, though Achilles walked ten times as fast as the tortoise. As soon as Achilles passes over the space separating him from the tortoise, the tortoise advances one-tenth of that space: Achilles passes over that tenth, but the tortoise has advanced a hundredth, and so on to infinity. This problem seemed to the ancients insoluble. The irrationality of the conclusion (that Achilles will never overtake the tortoise) arises from the arbitrary assumption of disconnected units of motion, when the motion both of Achilles and the tortoise was continuous."

4.3.2 The way forward, explains Tolstoy, is to introduce the dimension of time by making use of calculus, since:

> "This new branch of mathematics, unknown to the ancients, by assuming infinitely small quantities, that is, such as secure the chief condition of motion (absolute continuity), corrects the inevitable error which the human intellect cannot but make, when it considers disconnected units of motion instead of continuous motion."

4.4 *The Steam Engine*

4.4.1 One of the simple analogies cited by Tolstoy, the steam engine, is reminiscent of Adam Smith's observation that the 'invisible chain' linking two events may not be the true underlying causal mechanism:

> "Whenever I see a steam-engine move, I hear the whistle, I see the valve open and the wheels turn; but I have no right to conclude from that that the whistle and the turning of the wheels are the causes of the steam-engine's moving."

This analogy is expanded in the Epilogue, where Tolstoy identifies four stages of understanding as to how the steam-engine is moved:

Stage 1 *Pre-science*: "It is the devil moving it."
Stage 2 *Co-incident, but non-causal behaviour*: "The wheels are going round."
Stage 3 *First accompanying symptom*: "Smoke is being blown back from it."
Stage 4 *Ultimate cause*: "The steam compressed in the boiler."

4.4.2 Tolstoy extends this analogy to history as follows:

> "The only conception which can explain the movement of the steamer is the conception of a force equal to the movement that is seen.
> The only conception by means of which the movements of nations can be explained is a conception of a force equal to the whole movement of the nations."

4.5 *Ultimate Causes*

4.5.1 Tolstoy's final paragraph in Chapter 1 of Part XI is the quotation that appears at the beginning of my equities paper (1981):

> "For the investigation of the laws of history, we must completely change the subject of observations, must let kings and ministers and generals alone, and study the homogeneous, infinitesimal elements by which masses are led. No one can say how far it has been given to man to advance in that direction in understanding of the laws of history. But it is obvious that only in that direction lies any possibility of discovering historical laws; and that the human intellect has hitherto not devoted to that method of research one millionth part of the energy that historians have put into the description of the doings of various kings, ministers, and generals, and the exposition of their own views on those doings."

4.5.2 Both my price model for gilts (1978) and my price model for equities (1981) recognise that actual market prices will be the equivalent of 'arbitrarily selected units of motion', and hence discontinuous in the extreme, whereas the 'correct' prices should follow some smooth and continuous pattern in line with continuous variation in the underlying causal mechanisms. Also, I use calculus (two partial differential equations in each case) to identify a framework for 'central' prices about which market prices will fluctuate.

4.6 *Military Strategy*

4.6.1 In the next chapter, Tolstoy develops, in the historical context of the aftermath of the battle of Borodino, the theme that everything looks different with hindsight:

"The action of a commander-in-chief in the field has no sort of resemblance to the action we imagine to ourselves, sitting at our ease in our study, going over some campaign on the map with a certain given number of soldiers on each side, in a certain known locality, starting our plans from a given moment. The general is never in the position of the beginning of any event, from which we always contemplate the event."

4.6.2 Tolstoy then observes that apparently superb military plans generally fail to take into account crucial features of the circumstances of the moment, and illustrates this by showing that a retreat by the Russian army along the Kaluga road — an action that most historians said would have prevented the abandonment of Moscow to the French — ceased to be an option once Kutuzov gave a certain order not to join battle with the French:

"But the commander of an army has before him, especially at a difficult moment, not one, but dozens of plans. And each of those plans, based on the rules of strategy and tactics, contradicts all the rest. The commander's duty would, one would suppose, be merely to select one out of those plans; but even this he cannot do. Time and events will not wait."

4.6.3 Another serious constraint on the decision-making ability of the commander-in-chief is the degree to which his intellectual powers are distracted by a multitude of other necessary considerations, many of which have little or no bearing on the overall military strategy.

4.6.4 Furthermore, dangerously incomplete and inconsistent information means that important decisions will rarely correspond to 'rational behaviour' based on 'perfect information':

"The officer sent to inspect the locality comes back with a report utterly unlike that of the officer sent on the same commission just previously; and a spy, and a prisoner, and a general who has made a reconnaissance, all describe the position of the enemy's army quite differently."

4.7 *Kutuzov and Necessity*

Tolstoy concludes that Kutuzov's apparent freewill in being able to decide whether or not to abandon Moscow did not, in reality, exist, as an inevitable consequence of the types of constraints discussed earlier:

"Persons who forget, or fail to comprehend, those inevitable conditions under which a commander has to act, present to us, for instance, the position of the troops at Fili, and assume that the commander-in-chief was quite free on the 1st of September to decide the question whether to abandon or to defend Moscow, though, there could no longer be any questions on the subject."

4.8 *Napoleon and Necessity*

4.8.1 Tolstoy then turns his attention to Napoleon, the other central character of the action, and his period of occupation of Moscow. While Napoleon's general strategy and detailed planning were regarded as brilliant by most historians, Tolstoy observes that they simply did not work.

"But, strange to say, all these arrangements, these efforts and plans, which were no whit inferior to those that had been made on similar occasions before, never touched the root of the

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

matter; like the hands on the face of a clock, when detached from the mechanism, they turned aimlessly and arbitrarily, without catching the wheels."

4.8.2 Tolstoy then delves deeper into the manner in which Napoleon's strategy fell apart, and his commentary includes a number of areas where certain groups of individuals were motivated by undisciplined short-term greed with no rational thought for the adverse medium-term and long-term consequences:

"The establishment of a municipal council did not check pillage, and was no benefit to any one but the few persons, who were members of it, and were able on the pretext of preserving order to plunder Moscow on their own account, or to save their own property from being plundered.

The attempts to entertain the people and the troops with theatres were equally unsuccessful. The theatres set up in the Kremlin and Poznyakov's house were closed again immediately, because the actors and actresses were stripped of their belongings by the soldiers."

4.8.3 Very interestingly, Napoleon's belated decision to reverse his disastrous strategy was triggered, not by rational thought, but by precisely the same panic fear that Adam Smith refers to in his *History of Astronomy*, as only arising from the 'sudden apprehension of unexpected danger':

"The army, like a herd of cattle run wild, and trampling underfoot the fodder that might have saved them from starvation, was falling to pieces, and getting nearer to its ruin with every day it remained in Moscow. But it did not move. It only started running when it was seized by panic fear at the capture of a transport on the Smolensk road and the battle of Tarutino. The news of the battle of Tarutino reached Napoleon unexpectedly in the middle of a review."

4.9 *Pointers to a Better Framework*

In the second part of his Epilogue to *War and Peace*, Tolstoy philosophises at length on the inability of apparently omnipotent individuals to control events, and on the seemingly irreconcilable conflict between freewill and necessity. I have selected four areas as being of particular relevance in our quest for a better framework for human behaviour than the 'rational expectations' paradigm.

4.10 *Types of Upward Delegation*

Most historians of Tolstoy's day, or earlier, while defining power in general as "The combined will of the masses, transferred by their expressed or tacit consent to the rulers chosen by the masses," realised that, in practice, this upward delegation could be usefully classified into three types:

Type 1 — where the will of the masses is always unconditionally delegated over to that ruler or those rulers whom they have chosen.

Type 2 — where the delegation is under certain definite conditions.

Type 3 — where the delegation is conditional, but the conditions are uncertain and undefined.

4.11 *Uniting for Common Action*

An interesting 'law of human behaviour' is where Tolstoy observes:

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

> "For common action, men always unite in certain combinations, in which, in spite of the difference of the objects aimed at by common action, the relation between the men taking a part in the action always remains the same;"

and then observes that the higher an individual is ranked, within a military or other organisation, the less part that individual takes part in the day-to-day action:

> "The military organisation may be with perfect accuracy compared to the figure of a cone, the base of which, with the largest diameter, consists of privates; the next higher and smaller plane, of the lower officers; and so on up to the apex of the cone, which will be the commander-in-chief."

4.12 *Freewill and Necessity*

Tolstoy suggests that it is the hitherto unsolved conflict between freewill and necessity that has led to what he calls 'the false path that science has followed':

> "If one man only out of millions once in a thousand years had the power of acting freely, that is, as he chose, it is obvious that a single free act of that man in opposition to the laws governing human action would destroy the possibility of any laws whatever governing all humanity."

4.13 *Plastering over the Windows*

Finally, Tolstoy follows up his criticisms of conventional scientific methods with a very vivid and amusing simile:

> "The scientific men and their disciples who suppose they are solving this question are like plasterers set to plaster one side of a church wall, who, in the absence of the chief superintendent of their work, should in the excess of their zeal plaster over the windows, and the holy images, and the woodwork, and the scaffolding, and rejoice that from their plasterers' point of view everything was now so smooth and even."

4.14 *Freewill*

4.14.1 He then interprets 'freewill' in a very illuminating manner, namely as the unknown links in Adam Smith's 'invisible chain' of the true causal mechanisms that underly real world behaviour:

> "In the experimental sciences, what is known to us we call the laws of necessity; what is unknown to us we call vital force. Vital force is simply an expression for what remains unexplained by what we know of the essence of life. So in history what is known to us we call the laws of necessity; what is unknown, we call freewill. Freewill is for history simply an expression for what remains unexplained by the laws of men's life that we know."

4.14.2 We therefore find that the scientific viewpoints of Adam Smith and Tolstoy are remarkably similar. Smith not only regards it as the exception rather than the rule for us to discover the complete chain, but also recognises that, at any point in time, the then current state of scientific knowledge may be erroneously regarded by scientists, scholars and laymen alike as far more complete than it really is. Tolstoy concludes that much of the blame for the incompleteness of scientific knowledge must lie with the limited powers of the human intellect, and that, in the area of human behaviour, most of the difficulties are, so to speak, swept under the carpet by creating the fiction of individual freewill, which is often completely overwhelmed by the circumstances of events.

4.15 *The History of Astronomy*

In the final chapter of his Epilogue, Tolstoy — precisely as Adam Smith did more than a century earlier — uses the history of astronomy, and in particular the Copernican Revolution, as indicating the best way to make further progress in our understanding of human behaviour:

> "Just as then in the question of astronomy, now in the question of history, the whole difference of view rested on the recognition or non-recognition of an absolute unit as a measure of visible phenomena. For astronomy, this was the immobility of the earth; in history, the independence of personality — freewill.
>
> In the first case, we had to surmount the sensation of an unreal immobility in space, and to admit a motion we could not perceive of by sense. In the present case, it is as essential to surmount a consciousness of an unreal freedom and to recognise a dependence not perceived by our senses."

4.16 *Conclusions*

The parallel conclusions, to which I have been led by personal experience, are that the rational behaviour approach is seriously flawed as an explanation of human behaviour in the real world, and that we must find hitherto unrecognised dependences between actual behaviour and the circumstances of the situation before we can make any progress beyond our present, very incomplete, levels of understanding in social sciences such as economics.

5. Actuarial Science

5.1 *Introduction*

5.1.1 A very useful primary classification of the general stages of development of actuarial science (yet another example of the 'Adam Smith' approach to the philosophy of science) is given by Bühlmann (1987). This ASTIN Bulletin editorial article entitled 'Actuaries of the Third Kind?' traces the history of actuarial science from its birth in the 17th century, when it was exclusively devoted to problems of life assurance, to the present. 'Actuaries of the Second Kind' succeeded in having their methods applied to non-life insurance too, and 'Actuaries of the Third Kind' were now in the process of creating a new scientific philosophy for handling investment problems. Less than a decade later, we now have to add 'Actuaries of the Fourth Kind', who are doing pioneering work in the 'wider fields' arena.

5.1.2 This brief section asks two questions:

(1) What, if any, are the features that distinguish actuarial science from other branches of the social sciences?

(2) Can we have any confidence that actuarial science, particularly in the investment and 'wider fields' areas, will avoid the pitfall, described by Tolstoy, of producing spuriously elegant formulations that fail to address the fundamental issues?

5.2 *The Dimension of Time*

Perhaps the most fundamental model of actuarial science is compound interest, which provides a convenient framework for comparing financial values at different points in future time. Within financial economics, on the other hand, it has generally to be assumed (as in utility theory and the theory of games) that all outcomes are resolved instantly, or (as in the Sharpe Diagonal Model) that an arbitrary single period approach is appropriate.

5.3 *Actuarial Mathematics*

Whereas pure mathematics essentially involves the rigorous development of initially abstract concepts and axioms that may, or may not, have any real world interpretation, the main emphasis within actuarial mathematics is on finding, to a known degree of accuracy and using available numerical data, practical solutions to business problems involving uncertain future events. The training to use 'general reasoning' to justify a mathematical result is a further safeguard that, in actuarial science, mathematics is the servant, rather than the master, of theoretical development.

5.4 *Conclusions*

It is likely that the general methodologies and principles of actuarial science will provide the necessary numerical frameworks to develop relevant concepts of economic and financial behaviour into practical models, without having to introduce simplifying assumptions that might be inconsistent with the real world behaviour we wish to study and model.

6. Medical and Physical Analogies

6.1 *Introduction*

6.1.1 In this section I shall take further bearings, using 'superb' theories in the medical and physical sciences as reference points. I shall also comment on the role of a mathematical approach in deriving scientific theories that not only 'soothe our anxieties', but are also of relevance to practical action in the real world. I justify this seemingly unnecessary digression by citing two philosophers of the highest possible standing — Adam Smith and Erwin Schrödinger.

6.1.2 It seems likely that the major project that Smith had in mind, but was unable to complete within his lifetime, could be regarded as a much expanded version of his *History of Astronomy*, in which, by examining 'what came before' in advanced fields of human intellectual activity, we could find useful pointers to 'what might come after' in currently less advanced fields. The medical and physical sciences are, and have been since the time of the Ancient Greeks, far more advanced than the social sciences, of which economics is an important branch.

6.1.3 In the first chapter of his book *What Is Life*, Schrödinger justifies his incursion into fields outwith his primary fields of expertise by observing that the

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

vast explosion of human knowledge in recent times makes it virtually impossible for any one individual to realise that there may be relevant knowledge in other fields that can significantly enhance the rate of progress in his or her own fields of expertise.

6.2 *Bernoulli and Smallpox*

6.2.1 In 1760, some three years before the approach to statistical inference, which later became known as Bayes' Theorem, was communicated (after Bayes' death) to The Royal Society, Daniel Bernoulli — best known for his suggested resolution of the St Petersburg Paradox — wrote a paper entitled '*Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévénir*', in which he used 'inverse probability' to justify the conjecture that inoculation against smallpox, using material containing a mild strain of the organism causing the disease, was the probable cause of a significant reduction in the number of deaths as between those inoculated and those not inoculated.

6.2.2 Bernoulli's important contribution to the control of smallpox as a threat to human life was mathematical and statistical, rather than medical, in nature, and it set an example for other statistical applications in medicine and biology. To this day, clinical trials of new drugs are evaluated using similar principles, and it is interesting to note that many of the statistical tables in *Actuarial Tables for Examination Purposes* and in other scientific reference books and textbooks are reproduced from *Biometrika*.

6.3 *Pioneering Practical Work*

6.3.1 Bernoulli's work on smallpox, while not strictly medical in nature, was a valuable guide to leading figures in the medical world as to whether inoculation against smallpox and similar diseases was a rewarding field to pursue. Subsequent progress then depended on the resourcefulness, skill and ingenuity of pioneers on the practical side.

6.3.2 The crucial breakthrough in the case of smallpox was when Jenner found that inoculation with material containing the cowpox micro-organism gave total immunity from smallpox without, as sometimes happened when a weak strain of smallpox itself was used, leading to the death of the person inoculated. Only when the potentially fatal side-effects of inoculation had been thus eliminated could it be said that Europe had been liberated from one of its most serious scourges.

6.4 *Practical Judgements in the Real World*

After the refinement of medical techniques has been taken as far as possible, there remains the important question as to how best to apply these techniques in the real world in the light of all the ethical, economic and political considerations that have to be taken into account.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

6.5 *Causal Mechanisms*

6.5.1 Subsequent research work often reveals the true nature of what Adam Smith called the 'invisible chains' linking two initially disjoint events, and textbooks understandably concentrate on describing and classifying these causal mechanisms, rather than giving detailed accounts of previous (often statistical) far less complete states of understanding.

6.5.2 In the case of smallpox, Jenner's breakthrough is now viewed as a particular (albeit historically crucial) example of a certain type of immunological reaction. For example, in Keeton (1972), an authoritative biology textbook, the first paragraph under the heading of 'The Nature of Immunological Reactions' is as follows:

> "Immunological reactions have been the subject of much study since the English physician Edward Jenner discovered in 1796 that people develop immunity to smallpox if they are artificially injected with material that induces a very mild form of the disease (actually cowpox). Further dramatic demonstrations of the immune reaction were made by Louis Pasteur in France during the latter half of the nineteenth century."

6.5.3 No reference is made to Bernoulli's contribution, which helped pave the way for the previous methodology that was superseded, almost overnight, once Jenner published the results of his experiments. In general, textbooks must necessarily focus on a description of those practical techniques that are the best available at the time. Also, the more convincing the description of the true causal mechanisms that underly the 'state of the art' practical methodology, the closer we are likely to be to understanding the 'invisible chain' that leads to a highly satisfactory scientific system capable of guiding practical action.

6.6 *The Limitations of Mathematical Models*

6.6.1 While the great benefits of a mathematical approach in the development of economic theory are generally recognised, the possible dangers that can also arise receive very little attention. One of the most illuminating discussions of both the advantages and the disadvantages is given by Allais (1954). The author's English summary is a clear warning that the dangers were perceived to be very real:

> "It is the purpose of the following study to review as far as possible not only the immense advantages but also the possible dangers in the use of mathematics in economics.
>
> This study was written and submitted to *Econometrica* before Professor Morgenstern sent the Council of the Econometric Society his recent suggestions regarding the conditions that ought to be required for election to the Fellows of the Society and in which he emphasized the danger in economics of a purely abstract orientation. The present study is concerned with this same problem."

6.6.2 Towards the end of the paper, Allais stresses the important part that observation of the real world should play in economic science, and suggests that practical business experience is of far more relevance than book learning.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

6.6.3 Some 40 years later, Merton (1994) concludes his review of the application of mathematical models in finance practice with some words of caution about their use:

> "At times the mathematics of the models become too interesting and we lose sight of the models' ultimate purpose. The mathematics of the models are precise, but the models are not, being only approximations to the complex real world. Their accuracy as a useful approximation to that world varies considerably across time and place. The practitioner should therefore apply the models only tentatively, assessing their limitations carefully in each application."

6.7 *The Periodic Table in Chemistry*

6.7.1 In 1869 the Russian chemist Mendeleev discovered that elements could be arranged into rows and columns in such a manner that the elements in any one column have chemical and physical properties which are either similar or vary in a regular manner. Having noticed several vacancies in his table, Mendeleev not only predicted the discovery of new elements, but also used this periodic regularity to predict their chemical and physical characteristics.

6.7.2 One of the most important properties of an element is its metallic character, which represents its ability to lose its valence electrons to form the corresponding metal cation. Metallic character increases in moving along a row from right to left or in moving down a column. A highly schematic representation of a modern periodic table is given in Figure 6.1.

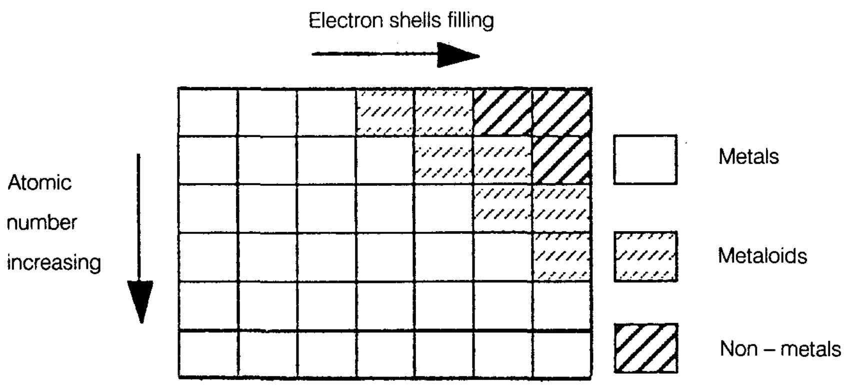


Figure 6.1.

6.7.3 I believe that theories within economic science can be similarly classified into rows and columns, in that the general power of a theory increases in moving from right to left along a row, and in moving down a column. Taking 'degree of uncertainty' increasing as the equivalent of electron shells filling, and 'level of understanding' as the equivalent of atomic number, gives a periodic table of the form shown in Figure 6.2.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

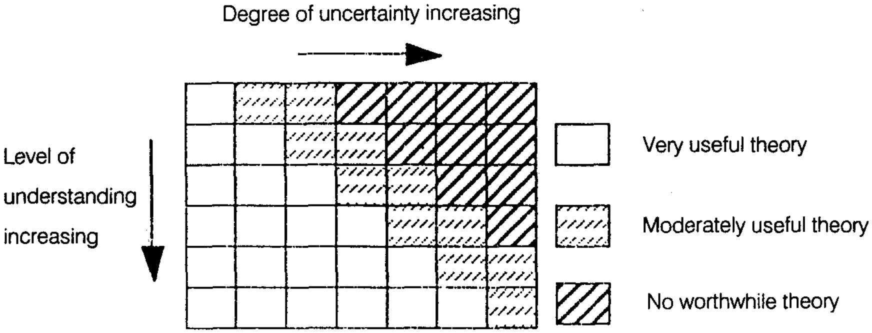


Figure 6.2.

6.7.4 I have introduced a strong shift to the left, as compared to the periodic table in chemistry, to reflect the obvious fact that human behaviour, and in particular human choice under uncertainty, is many orders of magnitude more complex than the physical behaviour of identical atoms.

6.8 *The History of Astronomy*

6.8.1 Adam Smith, Tolstoy and Kuhn (1970) have all suggested that successively better scientific explanations involve quantum leaps in 'levels of understanding' rather than the continuing refinement of any one paradigm of scientific thought. Since it is the same human intellect struggling to achieve order out of apparent chaos in both general areas, it seems not unreasonable to assume that the same quantum levels will apply in both the physical and the social sciences. We can then calibrate the scale for 'level of understanding', in the periodic table of economic theories, using astronomy as the benchmark. I now follow the narrative of the technical part of Smith's *History of Astronomy* to identify quantum levels of understanding in astronomy.

6.8.2 Smith, first of all, explains how the system of concentric spheres (with the earth at the centre) was "the first regular system of astronomy which the world beheld, as it was taught in the Italian School before Aristotle and his two contemporary philosophers Eudoxus and Callippus had given it all the perfection which it is capable of receiving."

6.8.3 While Eudoxus believed that 27 'celestial spheres' were adequate, other observers found it necessary to increase the complexity of the system to obtain a sufficiently accurate representation of real world observations:

> "Callippus, though somewhat younger, the contemporary of Eudoxus, found that even that number was not enough to connect together the vast variety of movements which he discovered in those bodies, and therefore increased it to 34. Aristotle, upon a yet more attentive observation found that even all these Spheres would not be sufficient, and therefore added 22 more, which increased their number to 56. Later observers discovered still new motions, and new inequalities, in the heavens."

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

6.8.4 The failure of this increasingly complex system to 'soothe the imagination' to a sufficient extent, led to the emergence of what is generally known as the Ptolemaic system:

> "Another system, for this reason, not long after the days of Aristotle, was invented by Apollonius, which was afterwards perfected by Hipparchus, and has since been delivered down to us by Ptolemy, the more artificial system of Eccentric Spheres and Epicycles."

6.8.5 Smith's description of the new paradigm as being 'more artificial' is most illuminating; it implies that no causal mechanism was put forward to justify the mathematical formulation that was used to 'explain' observed behaviour.

6.8.6 As more accurate measurements became possible over the centuries, the Ptolemaic framework had to be made more and more complex by the introduction of additional spheres to reflect these new observations. Various corrections were introduced in the tables produced by Almamon in the ninth century, but by the thirteenth century, even these tables were seen to be too inaccurate for practical use, and Alfonso X 'the Wise', king of Castile and Leon, gave orders for the compilation of more accurate astronomical tables. By the fifteenth century, these tables, too, were seen to be inadequate, and various philosophers began to search for a better framework than that of the, by then, horrendously complex Ptolemaic system.

6.8.7 Having examined all the hypotheses 'which the ancients had invented', and, in particular, having almost certainly been aware of the heliocentric hypothesis put forward by Aristarchos of Samos in the third century B.C., Copernicus realised that if the earth revolved each day around its axis, and if all planets (including the earth) revolved around the sun, then these two simple mechanisms "could, without the embarrassment of epicycles, connect together the apparent annual revolutions of the sun, and the direct, retrograde and stationary appearances of the planets."

6.8.8 Although Copernicus could see all the great advantages of his system over the Ptolemaic system, he was also well aware of the outright hostility that his view would attract. Having circulated his *Commentariolus* privately in 1514, he "began coolly to consider what a strange doctrine he was about to offer to the world, and he so much dreaded the prejudice of mankind against it that, by a species of continence, of all others the most difficult to a philosopher, he detained it in his closet for thirty years altogether."

6.8.9 The reception of his *De Revolutionibus*, when it eventually became public after his death, was precisely as he had feared:

> "When it appeared in the world, it was almost universally disapproved of, by the learned as well as by the ignorant. The natural prejudices of sense, confirmed by education, prevailed too much with both, to allow them to give it a fair examination. A few disciples only, whom he himself had instructed in his doctrine, received it with esteem and admiration."

6.8.10 To the sixteenth century equivalent of today's 'intelligent layman', there were three very reasonable grounds for rejecting the Copernican view of the heavens. First there were the 'natural prejudices of sense', in that the surface of the earth is perceived to be at rest, rather than hurtling through space at unimaginable speeds. Second, it was contrary to previous 'scientific' explanations that had been accepted both in institutions of higher learning and by religious authorities. Third, even astronomers were divided about the merits of the new system in the light of several perceived difficulties:

(i) There was no evidence that Mercury and Venus, the planets whose orbits should lie between the earth and the sun, exhibited phases similar to those of the moon.

(ii) There was no evidence of any other heavenly body orbiting a planet in the same manner that the moon was supposed to orbit the earth.

(iii) There was no evidence of any other heavenly body revolving around its own axis.

6.8.11 The general acceptance of the Copernican system as the new scientific paradigm was greatly speeded up by the work of Galileo, who was the first to realise that the recent invention of the telescope could be applied to great effect in astronomy. Despite the rudimentary nature of his telescope, he made three crucial discoveries:

(1) Mercury and Venus (as Copernicus had correctly predicted) did, indeed, exhibit phases, but these could not be detected by the naked eye.

(2) Jupiter had four satellites, which orbited the planet in precisely the same manner that the moon orbited the earth.

(3) There were sometimes 'spots' on the sun which moved over time in such a manner as to show that the sun revolved about its own axis.

6.8.12 In addition to resolving these technical astronomical difficulties, Galileo also addressed the first area of resistance by using various everyday examples (such as dropping a ball from the mast of a ship under sail) to explain why it is generally only relative motion, rather than absolute motion, that is perceived by our senses.

6.8.13 The next breakthrough was in the determination of empirical laws as a result of the combined work of Tycho Brahe, who achieved a far greater degree of accuracy of observations than earlier astronomers, and Kepler, who deduced somewhat laboriously from these observations that the orbits of planets around the sun were not, as Copernicus had assumed, perfectly circular in nature.

6.8.14 After stumbling down numerous blind alleys in misguided attempts to find 'harmony' in the orbits of the planets, Kepler eventually arrived at his three 'Laws of Planetary Motion':

First Law Planets follow elliptical orbits about the sun with the sun at one focus.

Second Law Planets sweep out equal areas of their elliptic orbits around the sun in equal times, and, in particular, move faster the closer they are to the sun.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

Third Law The squares of periodic times of planets around the sun are proportional to the cubes of their mean distances from the sun.

6.8.15 Smith describes how these Laws, while useful from a computational point of view, were almost beyond the grasp of the human intellect. Even the followers of Copernicus were reluctant to adopt "so strange a hypothesis" as the First Law. The Second Law was "too difficult to be followed", while the Third Law was "of too intricate a nature to facilitate very much the effort of the imagination in conceiving it."

6.8.16 The next quantum leap in understanding was Sir Isaac Newton's universal principle of gravity, which, in Smith's description, "made the most happy, and, we may now say, the greatest and most admirable improvement that was ever made in philosophy, when he discovered that he could join together the movements of the Planets by so familiar a principle of connection, which completely removed all the difficulties the imagination had hitherto felt in attending to them."

6.8.17 Newton realised that "if this attractive power of the Sun, like all other qualities which are diffused in rays from a center, diminished in the same proportion as the squares of the distances increased", then such an inverse squares law of gravitation would lead directly to all three of Kepler's Laws. Smith then describes the crucial supporting evidence from astronomical observations that allowed Newton to prove, beyond reasonable doubt, that gravity was indeed the true connection principle:

> "Experience shews us, what is the power of gravity near the surface of the Earth. That it is such as to make a body fall, in the first second of its descent, through about fifteen Parisian feet. The Moon is about sixty semidiameters of the Earth distant from its surface. If gravity, therefore, was supposed to diminish, as the squares of the distance increase, a body, at the Moon, would fall towards the Earth in a minute; that is, sixty seconds, through the same space, which it falls near its surface in one second. But the arch which the Moon describes in a minute, falls, by observation, about fifteen Parisian feet below the tangent drawn at the beginning of it."

6.8.18 My final level of understanding in astronomy relates to the prediction of the future appearances of comets. Smith's *History of Astronomy* refers to the first predicted return of Halley's Comet:

> "Comets had hitherto, of all the appearances in the Heavens, been the least attended to by Astronomers ..... Newton accordingly applied his mechanical principle of gravity to explain the motions of these bodies. That they described equal areas in equal times, had been discovered by the observations of some later Astronomers; and Newton endeavoured to show how from this principle, and those observations, the nature and positions of their several orbits might be ascertained, and their periodic times determined. His followers have, from his principles, ventured even to predict the returns of several of them, particularly of one which is to make its appearance in 1758. We must wait for that time before we can determine whether his philosophy corresponds as happily to this part of the system as to all the others."

6.8.19 We can continue the narrative almost a quarter of a millenium later, and with virtually no break in continuity, in Hughes (1987):

"Edmond Halley and his comet stand at the watershed of cometary science. Before his time, comets were regarded as evil omens, the precursors of doom, disaster, disease and the death of kings. Their physical nature, chemical composition and orbits were unknown. Isaac Newton showed the world how to calculate the orbit of a comet; Edmond Halley applied this technique to the data of twenty-four comets. He proved that one comet was periodic and as such was predictable."

6.8.20 Halley's crucial breakthrough involved the dimension of time; he made extensive use of data from the past, rather than relying on observations during his lifetime alone, to reduce the degree of uncertainty in his predictions of future behaviour. In 1726 Halley wrote that "It is probable that its return will not be until after the period of 76 years or more, about the end of the year 1758, or the beginning of the next."

6.8.21 Hughes records the actual outcome in the following terms:

"The comet actually passed perihelion (its closest point of approach to the Sun) on 13th March 1759, some 76.49 years after its previous passage. Astronomers were used to predicting planetary movements against a 'fixed' stellar background and now at least one comet could join that band of wanderers. People could look forward to its return whereas previously all comets crept up on them unannounced. The reputation of astronomy was considerably enhanced by the comet's return."

6.8.22 We can now summarise the 'levels of understanding' in astronomy by means of the following table:

Level	Characteristics	Key individuals
1	Simplest explanation	Eudoxus, Callippus, Aristotle
2	Provisional guiding principle	Apollonius, Hipparchus, Ptolemy
3	Improved guiding principle	Aristarchos, Copernicus, Galileo
4	Pioneering practical work	Tycho Brahe, Kepler
5	Causal mechanism	Newton
6	Advanced practical applications	Halley

6.9 *Human A-B-O Blood Types*

6.9.1 Four blood types are recognised, depending on whether or not cellular antigens A and B are present. The cells of type-A blood contain antigen A, the cells of type-B blood contain antigen B, the cells of type-O blood contain neither A nor B, and the cells of type-AB blood contain both A and B. The antigens A and B react with certain antibodies that can be present in the blood plasma. Normally a person's plasma contains antibodies corresponding to any antigens not present in the cells. However, if cells bearing a particular antigen and the plasma containing the corresponding antibody come together in a person's blood, the cells agglutinate (i.e. clump together) and death may result.

6.9.2 These properties have important implications for blood transfusions. As might have been expected, it is always best to obtain a donor who has the same blood type as the patient. However, it has been found that where such a donor is not available blood of another type may be used provided that the plasma of the patient and the cells of the donor are compatible. Ignoring the relatively rare type-AB, this gives the following table of relationships:

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

Blood Group	Can act as donor to
O	O, A, B
A	A
B	B

6.9.3 Let us now reconsider Keynes's observation:

> "For it is not sensible to pay 25 for an investment of which you believe the prospective yield to justify a value of 30, if you also believe that the market will value it at 20 three months hence."

We now postulate the existence of three distinct classes of investor. The 'intelligent' type-A investor, having realised that the price of 25 is below the expected long-term price of 30, but being unable to assess the likely short-term price behaviour, buys at 25. The 'unintelligent' type-B investor, aware of unfavourable short-term factors, but unable to establish whether the current price is high or low in relation to the long-term prospects, sells at 25 with no intention of buying back. The 'omniscient' type-O investor understands the psychology of both the type-A and the type-B investor, and calculates that the aggregate influence of type-A investors, type-B investors and other type-O investors will take the price to 20 in three months' time, after which the price will rise to its correct long-term value of 30 in twelve months' time. The type-O investor, therefore, sells at 25 and buys back in three months' time at 20.

6.9.4 If the price does, indeed, go to 20 after three months and then to 30 after twelve months, the outcomes for the various types of investor are as follows:

Investor type	Gain after 3 months	Gain after 12 months
O	5	$12\frac{1}{2}$
A	–5	5
B	5	–5

Over the short run of three months, the 'unintelligent' investor does as well as the 'omniscient' investor, while the 'intelligent' investor does badly over the period. Over the 'long run' of twelve months, however, the 'intelligent' investor does better than the 'unintelligent' investor, but not nearly as well as the 'omniscient' investor.

6.9.5 Assuming perfect foresight on the part of investors in those areas where they make assessments, the general characteristics of this hypothetical situation can be summarised as follows:

Investor type	Expected return % p.a.	Strategy	Investment policy
O	50	Dynamic	Sell at 25 and watch to buy back at 20
A	20	Static	Buy at 25
B	Negative	Hazy	Sell at 25

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

Each investor will believe that he or she is acting 'rationally', but their perceived returns and their investment policies are markedly different.

6.9.6 Keynes suggested that the key to successful investment was "anticipating the anticipations of others". The 'omniscient' investor can fully understand the psychology of both other types of investor, but neither the 'intelligent' investor nor the 'unintelligent' investor can fully understand the actions of any other type of investor. This gives precisely the same relationships table as for compatibility in the context of blood donors:

Investor type	Can anticipate the action of
O	O, A, B
A	A
B	B

6.9.7 The interactive system corresponding to this set of three investor stereotypes is my first guideline principle in the quest for a better framework for human behaviour in the face of risk and uncertainty. The postulates of 'rational behaviour' and 'stockmarket efficiency' correspond to the particular limiting case of this system where virtually all investors are of the 'omniscient' type and, as a result, the profit potential of identifiable anomalies is not sufficient to cover the transaction costs involved.

6.10 *Einstein's Special Theory of Relativity*

6.10.1 When 'classical' physics could not explain some unexpected empirical results, such as those of the famous Michelson-Morley experiment carried out in 1886, various theoretical physicists tried to construct new theories to 'explain' the apparently anomalous behaviour. However, the situation was not finally resolved until Einstein produced his Special Theory of Relativity, which expresses, in mathematical language, the fundamental principle that the laws of physics are independent of the frame of reference of the observer. A central feature of the theory is that the speed of light, which is constant to all observers, is a limiting value which cannot be exceeded.

6.10.2 Special relativity is, essentially, a statement of a consistency principle, and an illuminating everyday example, cited in Einstein (1920), involves dropping a stone out of the window of a moving train. From the frame of reference of the person dropping the stone, it falls — ignoring air resistance — in a vertical straight line until it hits the ground. From the frame of reference of an observer at rest on the embankment beside the railway line, the stone traces out a parabola. The physical event is the same in both cases, and common sense suggests that it should be possible to find a 'universal' description of the event with which both sets of observations are consistent.

6.10.3 My second guideline principle is the similar underlying consistency with which the human mind recognises each element of risk, whether physical or financial, as:

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

Risk = Probability x Impact

where 'probability' relates to the perceived likelihood of a particular adverse occurrence and 'impact' relates to the severity, in psychological terms, of the consequences. As with special relativity, there is an upper limiting value, namely the maximum level of risk that an individual is willing to accept. The existence of an upper limiting value is of little practical significance in the case of financial risk in certain areas such as portfolio selection. However, in the case of high risk areas such as futures and options, the upper limiting value of risk is of crucial importance.

6.11 *Einstein's General Theory of Relativity*

6.11.1 As a very sweeping simplification, it might be said that Einstein's General Theory of Relativity expresses, in mathematical language, the fact that physical behaviour (including the component of time) in a particular locality is not absolute, but depends, amongst other things, on the quantity of matter in that locality.

6.11.2 It is interesting to note that astronomy, because it had been for centuries one of the most advanced areas of applied mathematics, played a central role. Newtonian mechanics was unable to explain a very minor, but nevertheless very stubborn, anomaly in the precession of the observed orbits of Mercury. Einstein's ability to embed this previously inexplicable behaviour within the ambit of his new theory played an important part in its being seen as highly plausible by the scientific community. Also, the only other prediction of general relativity that had any possibility of being verified by the measurement techniques available at that time related to light beams being bent by gravitational fields. Although it is now generally accepted that the 1918 solar eclipse 'evidence' of light from stars being bent by the gravitational field of the sun by precisely the amount predicted by general relativity is inconclusive, in that experimental error was of the same order of magnitude as the effect being examined, the report of this famous international expedition of astronomers led to the almost overnight acceptance of the new theory.

6.11.3 Let us now consider the comments, towards the beginning of Bernoulli (1738), regarding the implications of a fixed monetary amount having to be assessed, not in absolute terms, but relative to the circumstances of the individual concerned:

> "Somehow a very poor fellow obtains a lottery ticket that will yield with equal probability either nothing or twenty thousand ducats. Will this man evaluate his chance of winning at ten thousand ducats? Would he not be ill-advised to sell this lottery ticket for nine thousand ducats? To me it seems that the answer is in the negative. On the other hand, I am inclined to believe that a rich man would be ill-advised to refuse to buy the lottery ticket for nine thousand ducats. If I am not wrong then it seems clear that all men cannot use the same rule to evaluate the gamble."

Bernoulli concludes that the value of an item must not be based on its price, but

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

rather on the 'utility' it yields. I would define utility, in this context, as "usefulness relative to the specific circumstances of the individual".

6.11.4 In broad conceptual terms, there is a strong link between Bernoulli's original framework for utility and Einstein's General Theory of Relativity. Both were new mathematical theories devised in response to the inability of earlier (and essentially linear) paradigms to explain highly specialised examples of anomalous behaviour. Bernoulli's original goal was to find a framework for human behaviour that would resolve the perplexing St Petersburg Paradox, while Einstein had in mind the inability of classical theory to explain certain isolated examples of physical behaviour.

6.11.5 My third guideline principle is a generalisation of this founding principle of what came to be known as utility theory, namely that a person's behaviour under conditions of risk and uncertainty is a function both of the present circumstances of the person and also of his or her past experience and training. The 'past experience and training' component is introduced to give explicit recognition to the obvious fact that human behaviour depends, amongst other things, on the knowledge base that the individual has built up either 'first hand' from direct experience, or 'second hand' from education and training.

6.12 *Conclusions*

Taken together, the three guideline principles described above constitute my conceptual framework for a much more detailed model of financial and economic behaviour than that provided by current theories. The next stage is to examine whether these broad principles are consistent with observations of the real world.

7. A Model for Gilt-Edged Securities

7.1 *Introduction*

In this section I discuss the broad concepts involved in the design and application of the gilt-edged model described in Clarkson (1978), with particular emphasis on the conclusions that can be drawn as regards patterns of real world financial behaviour and their likely causes.

7.2 *The Initial Insight*

My starting point was the realisation, in 1971, that the very general arbitrage condition that 'no blatant anomalies should exist' could be translated into a much more precise statement, namely that, for stocks of the same term to maturity, a switch from a medium coupon stock into a combination of a higher coupon stock and a lower coupon stock, to maintain running yield, should not result in a higher amount at redemption. This led very easily to a convexity condition, with linearity being a limiting case, whereas all previous approaches had assumed strict linearity of price with coupon.

7.3 *Empirical Tests*

It was trivial to show that gilt-edged prices, at that time, exhibited marked non-

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

linearity in the direction that I had expected. Accordingly, both the explanatory power (in terms of goodness of fit for the same number of variable parameters) and the potential profitability of a model incorporating this non-linearity would be higher than for existing models, which were linear in nature.

7.4 *Steps Towards a General Model*

Having obtained the most general solution to the second order system of partial differential inequalities, the most important steps thereafter were, first of all, to identify the general properties of three auxiliary functions and, then, to obtain computationally tractable formulations involving as few variable parameters as possible.

7.5 *Practical Considerations*

A delicate balance had to be struck between goodness of fit and statistical stability. Initially I used two variable parameters in the case of two of the auxiliary functions, and three in the case of the other. Following extensive back-testing, which threw up one instance of statistical instability in a period of particularly violent market changes, the number of variable parameters was reduced from three to two in the third auxiliary function.

7.6 *Results*

The general level of root mean square error was very low compared to existing linear models with the same number of variable parameters.

7.7 *Mean Absolute Deviation Analysis*

Given the somewhat arbitrary nature of the auxiliary functions, it was highly unlikely that deviations from the fitted model would provide the best available indicators of the cheapness or dearness of an individual stock. Mean Absolute Deviation analysis, as first introduced into the actuarial literature by Plymen & Prevett (1972), was, therefore, used to provide an assessment of short-term cheapness or dearness relative to the general market level. It was found that the multiplier of 1.6 used to set the upper and lower control limits for investment trusts also worked well for gilt-edged stocks.

7.8 *Curious Behaviour at the Micro Level*

After trying to use autoregressive integrated moving average (ARIMA) analysis to improve upon the Mean Absolute Deviation approach for assessing cheapness or dearness, I concluded that there was pronounced cyclicality which was incompatible with the essentially linear framework of ARIMA analysis. Further investigation showed that this cyclicality was caused by low coupon stocks rising to unsustainable relative ratings on market rises, and by high coupon stocks showing the inverse behaviour of falling to overly depressed relative ratings. This explanation was consistent with observed variations in the overall goodness of fit

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

of the model as measured by the root mean error, which rose appreciably on sharp market movements before declining again as market levels stabilised.

7.9 *Curious Behaviour at the Macro Level*

Early in 1977, after sharp falls in market yields which resulted in high coupon stocks standing significantly above their par values for the first time, the positive convexity which had been a feature for many years reduced to zero and then became negative, giving a price structure which was illogical and unstable on anything other than a very short-term basis, in that significant and virtually risk-free arbitrage profits could be made by switching from medium coupon stocks to a combination of high and low coupon stocks with equivalent terms to maturity.

7.10 *Market Structure*

A very useful description of the term structure of interest rates can be obtained very easily from the fitted model by calculating the par yield curve, which shows, for each value of term, the coupon of the stock that stands at par.

7.11 *Framework for Judgement*

Because of the complexity of the dynamics of price changes, the model is an information framework designed to assist difficult human judgements, rather than a 'black box' which prescribes a course of action that has to be followed slavishly.

7.12 *Previous Actuarial Work*

The model can be regarded as a non-linear generalisation of the approach described in Pepper (1964), which, like all yield curve approaches, results in a model where price is a linear function of coupon. Until the late 1960s the non-linearity present in the gilts market was not particularly pronounced, and, accordingly, linear models could give sufficiently accurate portrayals of market structure.

7.13 *Academic Work*

The apparently innocuous simplifying assumptions adopted result in a linear model and in yield curves which are either monotonic increasing or monotonic decreasing. During the 1970s, however, the actual market structure was distinctly non-linear, and the yield curve often had a maximum value somewhere between ten and fifteen years.

7.14 *Levels of Investor Behaviour*

Investors or investment analysts, who struggle to compare the rating of a stock relative to a price structure on some systematic basis, can be said to be following type-A, or 'higher level', behaviour. Unstructured qualitative judgements, such as 'buy low coupon (high volatility) stocks since the market is rising', correspond to type-B, or 'lower level', behaviour. Using mean absolute deviation analysis, or similar numerical comparisons, to determine when 'cheapness' or 'dearness' is close to an extreme value corresponds to type-O, or 'optimal', behaviour. In the light of a non-linear approach becoming far more powerful than the previous linear approach, we also need to introduce a further type-R, or 'rational', behaviour, which corresponds to the combination of type-O behaviour as regards general philosophy of approach and a continual endeavour to develop the best possible mathematical models.

7.15 *Systematic Irrationality and a Prediction*

It is often claimed by financial economists that any new approach which offers the promise of superior returns will very quickly be taken up by others, and thereby lead to the elimination of the opportunities it was designed to exploit. The development of gilt-edged models over the 15 years or so from 1972, when my non-linear approach was first made public at a seminar in London, tells a quite different story as regards this crucial dimension of time. The Bank of England, who had developed the very comprehensive 'rational expectations' model described in Burman & White (1972), very quickly modified this linear model to incorporate the non-linearity to which I had drawn attention. One stockbroking firm used my non-linear approach, but most others were very slow even to attempt to modify their existing linear models. By 1986, when around 30 gilt-edged market-making firms were set up in response to structural changes in the London Stock Exchange, some of these firms did not even have in place models of the previous linear methodology. In general, these firms soon incurred such serious losses that they had to cease trading. Systematic irrationality seems the most apt description of such an approach to capital market trading when potentially profitable models are available, and it seems safe to predict that similar, but even more pronounced, irrationality, will exist in other less mathematically tractable areas, with the consequence that structured approaches to investment will have a potential profitability far in excess of what is generally assumed.

7.16 *Conclusions*

Even in the case of fixed-interest securities, there is strong evidence to suggest that markedly different levels of behaviour exist, and that appropriate structured analyses can translate the price movements that result from these behavioural patterns into superior returns.

8. A Model for U.K. Equities

8.1 *Introduction*

8.1.1 Within a few months of moving to the investment division of a life assurance company, I realised that, as regards equity investment, the human brain was unable to comprehend even the current rating of a share on any consistent

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

basis without the assistance of some mathematical framework to provide a standardised frame of reference analogous to the yield curve for gilt-edged securities first described in Marshall (1953). In most stockbrokers' circulars, share ratings were described in terms of price-earnings (PE) ratios on both a historic (i.e. last reported full year) basis and on a prospective (i.e. next full year to be reported) basis, without any guidance as to what weight should be attached to each of these two measures.

8.1.2 Consistent with Tolstoy's comments on the necessity to introduce the dimension of time, I set up a computerised system of 'real-time' relative PE ratios, using geometric interpolation to reflect a constant inter-year 'force of earnings growth'.

8.1.3 While this standardised measure was a useful step forward, it was by no means obvious how it could be combined with a 'growth factor' to determine whether a share was 'cheap' or 'dear' in terms of its long-term growth prospects.

8.2 *The Initial Insight*

8.2.1 In May 1975 it occurred to me that a simple additive measure of four essentially independent growth components — turnover, margins, financial adequacy, and 'operational flexibility' — might offer a promising way forward. If each of these components was assessed on a prescribed numerical scale from 'exceptionally poor' to 'exceptionally good', then their sum might represent a satisfactory proxy for the highly elusive 'long-term growth' factor that most investors are endeavouring to assess.

8.2.2 Other things being equal, the higher the long-term growth measure the higher should be the relative PE ratio. A mapping of long-term growth against relative PE ratio should, therefore, after appropriate standardisation for 'other things which are not equal', provide a quantitative measure of 'cheapness' or 'dearness' against the general market level.

8.3 *Empirical Verification*

8.3.1 Pilot tests were carried out to assess whether the embryonic model described above had any predictive power as regards identifying 'cheap' or 'dear' shares. Since I attributed the (previously undetected) non-linearity in the gilt-edged market to heterogeneity of investors' circumstances in the area of tax, I expected an even stronger non-linearity with regard to income in the equity market, since some investors see immediate income as the priority while others see long-term capital growth as the primary objective. Other things being equal, the share price will, therefore, increase with the dividend payout ratio.

8.3.2 The other factor that I introduced at this preliminary stage was balance sheet strength. This was a crude 'risk adjustment' to the long-term growth measure (which tended to imply 'normal' economic circumstances) to reflect the fact that a weak balance sheet cannot only lead to corporate disaster in very harsh economic conditions, but also restricts the rate of growth that can be achieved in very favourable economic conditions.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

8.3.3 From my (as then) five years' practical experience of U.K. equity investment, my best guess was that, as a first approximation, the price-earnings relative of a share could be expressed as:

$$\text{Price-earnings relative (\%)} = 75 + \text{factor total}$$

where the values given to each of the three factors were based on the following scales:

Long-term growth rate	–20 (very low) to +50 (very high)
Dividend payout ratio	–15 (very low) to +15 (very high)
Balance sheet strength	–20 (very weak) to +15 (very strong).

Since it would have been unrealistic to expect these *ad hoc* rules to reproduce the market structure exactly, a process of least squares fit was applied, giving:

$$\text{Price-earnings relative (\%)} = A + B \text{ x (factor total)}$$

where A and B should be approximately equal to 75 and 1 respectively.

8.3.4 Around 100 shares were assessed in this way in June 1975, with three different investment analysts each covering about one third of the total. To test for any systematic bias in the way the analysts estimated the growth factor (the only qualitative one, as prescribed quantitative bases were used for the other two), a separate least squares fit was carried out for each of the three groups of shares. In each case a well-defined regression line resulted, with A and B closer to 75 and 1, respectively, than I had expected. The cheapness of each share was then assessed, using the ratio of its expected price-earnings relative (calculated from the appropriate analyst's regression line) to its actual price-earnings relative. (Using a single regression line would have been wholly inappropriate; the 'cheapness' element was likely to be smaller than the separation between the regression lines in most cases.) For each of the three groups of shares, a very strong correlation between this assessment of cheapness and the relative price performance was apparent within a few weeks, thereby demonstrating the utility of even a highly simplified version of the type of model I had in mind.

8.4 *Steps towards a General Model*

8.4.1 The formal derivation of my general model in Clarkson (1981) begins by expressing the price P of a share as $P(E, D, G)$, where E, D and G are the earnings per share, the dividend per share, and the long-term growth rate, respectively, and then noting that we can express this in terms of a function of only two variables:

$$P(E, D, G) = E.\,\mathrm{F}(R,G)$$

where R is the dividend payout ratio $\frac{D}{E}$.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

8.4.2 In my gilts model, the key to achieving such a good and stable fit with a small number of parameters was the introduction of the two orthogonal functions $h(n)$ and $f(n)$, each of which required only two variable parameters. In the case of the equities model, I show that it is reasonable to assume that, for a given change in either dividend payout ratio or growth rate, the effect on the share price is independent of the value of the other variable. This means that $F(R,G)$ can be expressed as the product of two functions each of a single variable:

$$F\,(R,G) = F_1(R).F_2(G).$$

This elegant separation of the variables again results in two orthogonal functions, whose general properties can be determined using common sense investment principles.

8.4.3 Since the scale for the long-term growth rate factor G is arbitrary, we can, in principle, use an elementary compound interest formulation:

$$F_2(G) = (1 + g)^G$$

where g is a positive variable parameter.

8.4.4 In the case of $F_1(R)$ we can assume that $F_1(1) = 1$, i.e. the function has unit value when all available earnings are paid out in dividends. The interesting value is then $F_1(0)$, i.e. the value when it is company policy to distribute no earnings by way of dividends. Since companies of this kind exist and have a non-zero share price, $F_1(0)$ will be positive. Also, because of the income preference of some investors, as discussed above, $F_1(0)$ will be less than 1. The formulation I chose was:

$$F_1(R) = r + (1-r)e^{-d\left(\frac{1}{R}-1\right)}$$

where $0 < r < 1$ and d is a fixed positive number in the range 0.5 to 2.
I suspect that the much simpler linear formulation:

$$F_1(R) = r + R - rR$$

would have been just as good for all practical purposes.

8.4.5 Since many U.K. companies have a significant proportion of earnings arising from operations abroad or from exports, the long-term growth rankings would, in theory, have to be recalculated every time that it was thought that there had been a change in the relative economic merits of the relevant overseas countries as against the U.K. To avoid this potentially serious problem, I introduced an additional multiplicative factor:

$$F_3(A) = (1 + a\,A)$$

where A is the proportion of earnings arising abroad and a is a variable parameter which can be either positive or negative.

8.4.6 As in the case of the pilot model, a term for balance sheet strength is

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

required. This is introduced as a multiplicative factor $F_4(B)$, where B is the level of borrowings and $F_4(B)$ decreases as B increases.

8.4.7 Finally, a scaling factor k has to be introduced, giving the general model for the expected price P as:

$$P = k.\,E.\,F_1(R).\,F_2(G).\,F_3(A).\,F_4(B)$$

where the various component functions are as described above.

8.5 *Practical Considerations*

8.5.1 Since the model is, of course, little more than a structured framework for allowing investment analysts to translate their extensive, but generally unstructured, knowledge about a company into a single ranking measure, which can be compared with the market rating of the company's shares, by far the most important practical consideration is the preparation of suitable guidelines for the long-term growth rate. In the case of the gilts model, a delicate balance had to be struck between goodness of fit and statistical stability. In the equities case, a very difficult balance had to be struck between being too prescriptive, which would have stifled analyst flair and led to an expensive set of precise, but somewhat useless, numbers, and being too vague, which would have resulted in little or no consistency being achieved and to an expensive set of imprecise and totally useless numbers.

8.5.2 The 'middle ground' guidelines that I adopted were essentially as follows:

> "Regard the long-term growth rate of a company as being determined by four largely independent components:
>
> (1) rate of growth of turnover;
> (2) profit margins;
> (3) ability to finance expansion; and
> (4) operational flexibility.
>
> In assessing turnover, different companies within the same sector will vary greatly in the degree to which they anticipate and exploit potentially profitable markets. An assessment of these management qualities must be incorporated in the turnover components for each company. As regards profit margins, both the general level and the stability of margins have to be taken into account. Price sensitivity, the importance of the product, possible sources of competition, and price controls (formal or informal) are some of the many considerations involved. In assessing a company's ability to finance expansion, the main considerations are the general capital requirements of the principal activities, the level of retained earnings, the capital structure, and the amount of unused borrowing facilities. The fourth component, operational flexibility, takes into account not only the extent to which a company can adapt to changing economic circumstances but also any other relevant factors not taken into account in the other three components. Once you have researched each component, give it a numerical value on a scale of 1 (very poor) to 9 (very good) using non-integral values (such as $4\frac{1}{2}$ or $3\frac{3}{4}$) whenever you believe that integral values do not provide a fine enough scale Finally, your long-term growth measure G for each company is the sum of the your four component values."

8.5.3 Another very difficult balance that has to be struck relates to the

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

proportion of companies that can be accommodated within the general model without any *ad hoc* adjustments. The life assurance industry's philosophy of accepting the vast majority of proposers at 'normal rates' suggests that the optimal balance between commercial success and theoretical precision might be to include as many companies as possible within the framework of the general model, with special adjustments being necessary only in very exceptional circumstances.

8.5.4 Two examples of 'very exceptional circumstances' are where a take-over bid or a significant disposal of assets is likely. As an illustration of how such non-standard situations can be taken into account, the latter 'de-merger' scenario is described as follows in Clarkson (1981):

> "Consider the case of a company which decides to dispose of certain assets by forming a new company to hold these assets and then issuing shares in this new company to existing shareholders free of payment. Suppose that on the basis of the price model, taking into account earnings and dividends only, the share price would have been 100p, that the relevant asset value per share is 20p, and that the new shares are expected to stand at a 25% discount to net asset value. The additional value per share attributable to the disposal of assets is then 15p, and a multiplicative factor of 1.15 must be introduced to take the expected price to 115p ...... Where it was likely, but not certain, that such a disposal of assets would be made, a smaller factor, 1.1 say, might be appropriate to allow for the probability involved."

8.5.5 Another crucial consideration was the number of variable parameters to be used in the fitting process. In view of the exceptionally high 'noise' content of all equity markets, there was a serious risk that even three variable parameters could have led to too flexible a fit, in which case apparent changes over time, in the cheapness or dearness of shares, would have been statistical illusions, rather than a reflection of the underlying reality of the situation. Accordingly, only the growth rate parameter was allowed to vary during the first three years of operating the model.

8.5.6 As with the gilts model, the appropriate fitting process is to minimise the sums of the squares of the proportionate price errors. The computational aspects presented no problems, in that a well-defined minimum value for the sum of the squares of the proportionate price errors could be found very easily by elementary search procedures.

8.6 *Results*

8.6.1 The model was first run on 28 October 1975, with the universe of 136 shares being split into five virtually equal groups of 'buys', 'buy/holds', 'holds', 'hold/sells' and 'sells', based on the proportionate price error, which was interpreted as a measure of cheapness or dearness relative to the market. Since virtually no shares with very high growth rates G, were in the 'buy' group and virtually no shares with very low growth rates were in the 'sell' group, the original (and highly arbitrary) scale used for G was modified to eliminate this serious bias.

8.6.2 The average performance of the 'buy' and 'sell' groups, relative to the average for the universe of shares covered, is shown at fortnightly intervals in the following table. The 'average' figure is the average of the outperformance of the 'buys' and the underperformance of the 'sells'.

Period in weeks	Buys	Sells	Average	Trend
	%	%	%	%
2	+2.1	−1.1	1.6	0.6
4	+2.4	−3.4	2.9	1.2
6	+2.3	−2.1	2.2	1.6
8	+2.7	−2.7	2.7	2.1
10	+3.4	−2.7	3.1	2.5
12	+4.4	−1.3	2.8	2.8
14	+4.6	−0.4	2.5	3.2
16	+5.0	−0.8	2.9	3.4
18	+5.6	−2.0	3.8	3.7
20	+4.5	−2.6	3.6	3.9
22	+4.5	−2.2	3.4	4.1
24	+7.8	−2.3	4.0	4.3
26	+7.0	−3.2	5.1	4.5

Since the general pattern that emerged from 28 October 1975, and later, runs of the model was an average which seemed to be trending to a limit of around 6%, with half of this achieved within around 3 months, and around three-quarters in 6 months, a final 'trend' column is included using the formula:

$$\text{Average performance at time } t \text{ (in years)} = 0.06\,(1 - 0.0625^t).$$

8.6.3 To show the impact of this trend rate of outperformance after allowing for transaction costs, suppose that an initial portfolio of all the 'buys' is switched at predetermined fixed intervals, with all shares not then being 'buys' (assumed, for simplicity, to be 80% of the portfolio) being sold and re-invested in shares that are 'buys'. The effective annualised rate of outperformance or underperformance varies both with the switching period and with the 'round trip' switching expenses as shown below:

Switching period (in months)	Annualised relative performance			
	2% costs	3% costs	4% costs	5% costs
	%	%	%	%
1	−4.5	−13.4	−21.5	−29.0
2	3.6	−1.4	−6.1	−10.7
3	5.5	2.1	−2.3	−4.4
4	6.0	3.4	0.1	−1.6
5	6.0	3.9	1.9	0.1
6	5.7	4.0	2.3	0.6
7	5.4	4.0	2.5	1.1
8	5.1	3.8	2.6	1.3
9	4.8	3.6	2.5	1.4
10	4.5	3.5	2.4	1.4
11	4.2	3.3	2.3	1.4
12	3.9	3.1	2.2	1.4

For an institutional investor paying 2% costs, the optimal switching period of around 4 months gives a highly satisfactory outperformance of 6% p.a., whereas a private individual paying 5% costs (e.g. $1\frac{1}{2}$% stockbrokers' commission each side, $1\frac{1}{2}$% bid-offer spread, and $\frac{1}{2}$% transfer duty) would see most of the potential gains swallowed up in dealing expenses, even at the optimal switching period of around 10 months.

8.6.4 When all three variable parameters were introduced, with appropriate safeguards, in September 1978, it was found that their values were reassuringly stable within sensible ranges, and that they moved over time in a manner that accurately reflected aggregate investor sentiment in the market. In particular, the dividend parameter r, which was assigned the value 0.6 until September 1978, varied between 0.48 and 0.62 during 1979 and 1980, thereby confirming the validity of the reasoning in ¶8.4.4.

8.7 *Mean Absolute Deviation Analysis*

8.7.1 Given the practical difficulties of obtaining consistent forecasts of the future earnings profiles of different companies, it is highly unlikely that the values of the growth rates G, used in the model, will be in close agreement with the 'consensus' values implicit in market prices. Mean Absolute Deviation analysis, as first described in the actuarial literature by Plymen & Prevett (1972), is therefore used to supplement the essentially very long-term cheapness measure represented by the proportionate price error. As with gilts, this involves calculating, first, the geometric moving average of the price residual, then the absolute deviation, as the geometric moving average of the deviation from this central value, and finally setting upper and lower control limits at fixed multiples of this absolute deviation above and below the central value. The interesting question was whether the same multiplier of 1.6, which worked so well for investment trusts and gilts, would also be satisfactory for equities in general.

8.7.2 As shown, in particular by the detailed example involving Whitbread shares, it was found that the same value of 1.6 for this multiplier worked very well. The crucial significance of this 'investment constant' of 1.6 is discussed in detail in a later section.

8.8 *Curious Behaviour at the Macro Level*

8.8.1 In August 1975, when long-term growth rates were being estimated for the first time, particular regard was paid to the dangerously high rate of inflation in the U.K. It appeared that many capital-intensive companies, including most engineering companies, would be unable to generate sufficient finance internally to support the rate of profits growth previously experienced. This was confirmed by the fairly heavy rights issues (typically 1 for 3, at only a modest discount to the market price) being announced by many engineering companies. In view of this significant downgrading of the long-term growth prospects, most of the engineering shares in the main U.K. equity portfolio were sold, and the proceeds reinvested in companies expected to show much higher earnings growth in an environment of high inflation.

8.8.2 Over the ensuing six months, the value of the shares sold soared relative to the market by nearly 40% on 'recovery hopes', while the value of the shares purchased rose only modestly in relative terms, giving a seemingly disastrous loss on the switch operation. Thereafter, however, the sales began to underperform, while the purchases began to outperform. 'Break even' was achieved by the middle of 1977, and by the autumn of 1980, some five years after the switch was carried out, the purchases were around three and a half times the value of the sales.

8.8.3 The general lesson, as regards the price formation process in an equity market, is that short-term prospects (in this case the likelihood of a cyclical recovery in 1976) are often the primary focus of 'consensus' thinking, to the exclusion of any reasoned appraisal of the longer-term fundamentals.

8.8.4 The teachings of Modern Portfolio Theory (MPT) suggest that, in a falling market shares with a high 'beta', which will correspond broadly to shares with a high value of my long-term growth rate, will underperform as a class, and thereby lead to a decrease in my growth rate parameter g. This was not what happened between May 1979 and November 1979, when the All-Share Index fell by more than 20%, but the growth rate parameter g showed an upward trend. Closer examination of the fundamentals of the situation showed that the market weakness had been caused by falls in the prices of low growth companies, where both the profits and the cash flow had come under severe pressure in the harsh new economic environment, while the profits of high growth companies were, in general, only marginally affected. Selling 'high beta' shares to obtain some protection from this falling market would have been quite the wrong response.

8.9 *Curious Behaviour at the Micro Level*

8.9.1 It might be thought that differing investment recommendations regarding the same company will relate primarily to differences in the estimates of future profits and dividends. Again using Whitbread as an example, I show, in Clarkson (1981), that diametrically different recommendations can arise from precisely the same set of estimates.

8.9.2 A typical 'buy' recommendation in June 1980 was:

> "Although Whitbread shares have had a strong performance recently, in anticipation of these good results, the outlook is still favourable, and they are attractive as the leading major in a sector which has growth prospects in the coming year, despite the current difficulties."

The two 'sell' recommendations (out of 16 recommendations in total) were:

> "Whilst the trading/financial strength of the company means that it will weather the recession better than industry in general and the brewing industry in particular, a prospective fully taxed price-earnings ratio at a 48% premium to the market has more than discounted these positive factors,"

and

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

"The shares have outperformed the market by 28% over the past 12 months, and now command a massive premium in both price-earnings relative and yield terms, which more than takes account of the group's good performance under inflation accounting and the possibility of slightly above average profits growth this year."

8.9.3 The 'buy' recommendation had no numerical framework against which to make a reasoned case for the likely relative price behaviour, whereas both the 'sell' recommendations used a numerical frame of reference not dissimilar to my model, which, at that time, was also suggesting that Whitbread shares were very overvalued.

8.9.4 This example is one of many that suggested to me that investors who use any disciplined numerical approach for the assessment of 'cheap' or 'dear' shares are very much in the minority, in which case material inefficiencies in the pricing process can be expected to be the rule rather than the exception.

8.10 *Resultant Structure*

8.10.1 It will be obvious from the various aspects of 'curious behaviour', discussed above, that the price formation process in the U.K. equity market bears little or no resemblance to the simplistic model used in most academic work, namely that intense competition between expert investors leads to an equilibrium position, where (at least as a first approximation) the risk-adjusted returns on all shares are equal, and where all price changes are due to the arrival of new information.

8.10.2 Although the resultant structure is highly complex, with a very high level of background 'noise', practical experience of operating the equity model suggests that the price formation process, relative to a broad market index, involves three quite distinct 'forces'. The 'strong' force relates to fundamental cheapness, as assessed by comparisons across the whole market of long-term growth against current rating. The 'semi-strong' force, which exhibits marked cyclicality with 'white noise' superimposed, is the variation over time in the fundamental cheapness; mean absolute deviation charts are a very useful analytical tool. The 'weak' force relates to changes over time in aggregate investor preferences; such changes are very difficult to predict in advance, but their effects can often be observed in terms of the behaviour of the fitted parameters.

8.10.3 This gives the following summary table:

Price formation force	Nature	Predictability
Strong	Fundamental cheapness	Moderate
Semi-strong	Changes over time in fundamental cheapness	Moderate
Weak	Changes over time in aggregate investor preferences	Poor

8.11 *Framework for Judgement*

8.11.1 A commendably concise description of my general framework for

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

equity investment was given by Bishop in the discussion at the Institute (see Clarkson, 1981):

> "The paper is an attempt to systematize by mathematics a process which is akin to that engaged in by many professional investors in equities. This is agreeably unlike MPT, which consists of an enormous body of academic theorizing that only rarely makes any significant contact with the practical world."

8.11.2 The model is, indeed, a mathematical framework which can be used to enhance, but never replace, the judgement of skilled investment professionals. In particular, there are four quite separate ways in which value is added through a systematic approach:

(1) ensuring that a continuously updated record of individual share ratings relative to the market is maintained;

(2) assessing long-term growth prospects in far more disciplined a fashion than might otherwise be the case;

(3) comparing growth prospects and relative ratings across the entire universe of shares under consideration in a systematic manner that is totally beyond the capability of the unaided human brain; and

(4) comparing changes in cheapness or dearness over time in a systematic manner that is, again, totally beyond the capability of the unaided human brain.

8.12 *Earlier Actuarial Work*

8.12.1 One of the earliest types of equity model discussed in the actuarial literature is the dividend valuation model, in which the expected value of a share is equated to the present value of future dividends. As well as numerous practical difficulties (e.g. choice of discount rate), there is an even more serious conceptual problem, namely that, as the dividend payout ratio decreases to zero, the share price also decreases to zero, which is not what happens in the real world.

8.12.2 There are numerous similarities between my equity model and the linear regression model described in Weaver & Hall (1967), but also three very important differences. Firstly, in my model the dividend payout ratio function tends to a value of around 0.55 as the dividend payout ratio tends to zero, whereas in the Weaver & Hall model the corresponding limiting value is zero, leading to a serious bias (as commented on by several speakers in the discussion) in favour of high dividend payout (and generally low growth) companies. Secondly, the Weaver & Hall model uses the historic growth rate explicity; I use a wholly prospective basis while accepting that this will lead to a poorer statistical fit. Finally, the Weaver & Hall model uses no fewer than four variable parameters for the 'growth' function, whereas I use only one. My suspicion is that these four parameters, will lead to too flexible a fit, in that erratic price behaviour would tend to be absorbed by changes of parameters, rather than being flagged as exploitable anomalies.

8.12.3 Hempsted (1961) describes an alternative approach for the estimation of 'expected yields', using, in the first instance, reinvestment returns calculated from company accounts.

8.13 *Academic Work*

8.13.1 The experience obtained from operating the model as a practical investment tool suggests that many of the key principles of financial economics, while useful as a first attempt to bring capital market behaviour within a scientific framework, are inconsistent with the real world. The most striking inconsistency is in the area of so-called 'stockmarket efficiency', where the general academic viewpoint is that, other than in exceptional cases, it is futile to attempt to outperform the market on a risk-adjusted basis. The paradox of why such an obviously illogical conclusion became a central pillar of modern finance theory is discussed in a later section.

8.13.2 If the Miller-Modigliani 'dividend irrelevance proposition', that changes in the dividend payout ratio have no effect on a company's shares, was true in the U.K. equity market, then the dividend parameter *r* in my model would be equal to 1, whereas, in practice, a value of around 0.5 to 0.6 is obtained, confirming my prior expectation, as discussed in ¶8.4.4, that the obvious preference of many investors for immediate income will have a material impact on the general market structure. Prior to modern finance theory, the 'standard view' in the United States of America, as expounded by Graham & Dodd (1951), Williams (1938) and many others, was, again, that low-payout companies will sell at a discount to companies with an average payout ratio.

8.13.3 Miller & Modigliani (1961) reject this standard view in the following terms:

> "If such indeed were the case — and we, at least, are not prepared to concede that this has been established — then the analysis presented in this paper suggests there would be only one way to account for it, namely systematic irrationality on the part of the investing public."

8.13.4 My proposition is that their entire analysis is unsound, since one of their basic assumptions relates to 'rational behaviour', and is inconsistent with all the real world empirical evidence:

> "Rational behaviour means that investors always prefer more wealth to less and are indifferent as to whether a given increment to their wealth takes the form of cash payments or an increase in the market value of their holdings of shares."

8.13.5 The Sharpe Diagonal Model, in which 'beta' became the primary characteristic of each share, was undoubtedly an important theoretical breakthrough, in that it provided the first computationally feasible methodology for implementing the previously impracticable mean-variance framework described by Markowitz (1952, 1959). However, as described in ¶8.8.4, simplistic investment policies based on the Sharpe Diagonal Model can worsen, rather than enhance, portfolio performance.

8.14 *Levels of Investor Behaviour*

The same general patterns as for gilts emerge. Investors, or investment analysts, who struggle to compare the current rating of a company against long-term prospects on some systematic basis, are following type-A, or 'higher level', behaviour. Unstructured qualitative judgements, such as the 'buy' recommendation in ¶8.9.2, correspond to type-B, or 'lower level', behaviour. Using mean absolute deviation analysis or similar numerical comparisons to help determine when 'cheapness' or 'dearness' is close to an extreme value, corresponds to type-O, or 'optimal', behaviour. Type-R, or 'rational', behaviour corresponds to the combination of type-O behaviour as regards general philosophy of approach, and a continual process of self-assessment and self-improvement to assess, and, where practicable, to improve the techniques and skills required at each stage in the process.

8.15 *Systematic Irrationality and a Prediction*

8.15.1 In the final paragraph of Clarkson (1981), before the acknowledgements section, I observe that investment management is a highly competitive activity, and that any new approach which offers the possibility of superior performance will either flourish or be ignored, depending on the investment community's appraisal of its merits as compared with existing approaches. From my purely personal viewpoint, I believe that the failure of many in the academic community to accommodate, and then, where appropriate, to enhance structured fundamental approaches of the type I describe, can be classified as 'systematic irrationality'; it not only prevents financial economists from acquiring relevant practical skills, but — more seriously — it hinders the rate at which investors of all types will improve their competence when the conventional scientific wisdom is that no such improvements are feasible.

8.15.2 My prediction is that the swing away from MPT and related methodologies, when it eventually occurs, will be swift, irreversible, and driven by perceived commercial advantage.

8.16 *Conclusions*

My general conclusions are twofold. First, the evidence in support of the existence of different levels of behaviour is, I believe, overwhelming. Second, the teachings of financial economics, particularly in the areas of stockmarket efficiency and the Sharpe Diagonal Model, have hindered, very significantly, the development of practical approaches which can translate the observed dynamics of price movements into superior returns.

9. Improving the Performance of Equity Portfolios

9.1 *Introduction*

Clarkson & Plymen (1988) was written to draw attention to what we perceived to be the very limited practical relevance of MPT.

9.2 *General Comments on Modern Portfolio Theory*

9.2.1 While proponents of MPT claimed that it was the only practical

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

approach to investment management that incorporated a scientific framework for risk, we had grave general reservations, many of which are mirrored in the following comments made by Grimes in opening the discussion on Moore (1972):

> "The Markowitz 'minimum risk' portfolio was only a portfolio which had a minimum risk if the data were precisely correct, which was very unlikely. Among all the portfolios that a fund might have in practice, there was one that actually had minimum risk, and there was another, probably different, which inaccurate estimates and the Markowitz theory would imply had minimum risk, but all of those practical portfolios had risks so close to the minimum that there was no reason to choose between them. In portfolio selection, common sense in spreading investments was all that was required to achieve the amount of efficiency in avoiding risk."

9.2.2 The crucial point here is that scarce manpower and computing resources, which could probably be better utilised in fundamental analysis to improve the performance of the fund, will be diverted to elaborate, but spurious, computations that are of little, or no, benefit to the fund.

9.2.3 Certain aspects of the practical implementation of the equity model described in the previous section demonstrate the scale of the difficulties caused by estimation errors. The pilot tests showed how different analysts, given the same detailed guidelines, used quite different assessment scales for their numerical estimates. An even more serious problem relates to the final numerical output of even this very elaborate model being a ranking of attractiveness, rather than a direct forecast of the expected price change over any period. For instance, while the average cheapness of the 'buys' might have been around 15%, it would have been naive in the extreme to interpret this as implying 15% outperformance over some future period.

9.2.4 The seriousness of these estimation and standardisation difficulties, in the context of the Sharpe Diagonal Model, is confirmed by Hodges & Brealey (1971):

> "The analysis so far has assumed that we could effectively eliminate any exaggeration from the expected return and variances. We have argued that in practice most investors tend to make extravagant claims as to their forecasting abilities and find it very difficult to provide the expected values. We suspect that users of portfolio selection models have seldom made any serious attempt to overcome this problem, but have simply relied on forecasts as given by the analyst."

9.2.5 A quite separate problem in applying portfolio selection techniques is that risk may be double-counted. Pegler (1948), for example, defined investment policy as 'maximising the expected yield'. Because of various margins introduced into the detailed estimates, this yield was effectively net of risk. Similarly, since the detailed equity models described in Weaver & Hall (1967) and Clarkson (1981) both incorporate features which specifically penalise earnings volatility, the final rankings of attractiveness can be regarded as already containing an allowance for risk.

9.3 *Stockmarket Efficiency*

9.3.1 One might have expected portfolio theory to have progressed simultaneously on two fronts, firstly seeking out improved methods of fundamental analysis to identify cheap and dear shares, and, secondly, balancing risk on some commonsense basis against the potential relative outperformance over some selected benchmark. However, largely as a result of the empirical work carried out by Fama (1970) and Jensen (1968), the Efficient Market Hypothesis became a central feature of the teachings of financial economics, with the result that portfolio theory became more and more dominated by the notion that, since nothing useful could be done to improve the expected return, the entire focus had to be on the risk front.

9.3.2 In Clarkson & Plymen (1988) we explain, in some detail, why we reject the so-called scientific proof of efficiency. Our main argument is that the Sharpe Diagonal Model is such an inaccurate reflection of real world behaviour that it was unscientific in the extreme even to attempt to use it as a frame of reference against which to measure specific elements of real world behaviour.

9.4 *Fundamental Analysis*

9.4.1 We describe the philosophy, general structure, and practical results of the equity model discussed in the previous section. This model is put forward, not as the unique solution to the problem, but as an example of the type of systematic approach to fundamental analysis that is likely to enhance the performance of an already competent equity investment operation.

9.4.2 Even well-documented models of this type tend to be ignored by financial economists, no doubt as a result of the pervasive influence of the so-called scientific evidence of stockmarket efficiency. While an impartial comparative review of the strengths and weaknesses of the equity models described in Hempsted (1961), Weaver & Hall (1967), and Clarkson (1981), would be of considerable interest to those involved in the management or supervision of an equity investment operation, and might point the way towards the construction of a better model that incorporated some of the best features of each existing model, no such review is likely to be attempted by academic researchers, since their training leads them to believe that the opportunities which the models are designed to exploit do not exist in the first place.

9.5 *A Turing Test for Equity Models*

9.5.1 Penrose (1989) describes the 'Turing test', suggested by the British mathematician Alan Turing, as a systematic procedure for determining whether a computer could reasonably be said to think. I describe below my suggestion for a corresponding 'Turing test' for equity models, to resolve, once and for all, the question of whether a model can achieve consistent above average performance.

9.5.2 For a universe of at least 100 shares (to ensure that the average performance figures are not unduly distorted by mere random chance), the model is used to identify the 20% of shares which are most attractive (the 'buys') and the 20% of shares which are the least attractive (the 'sells'). At weekly intervals, from the selection date, calculate the average performance relative to the performance of the universe of shares for:

—M(B): the 'buys' selected by the model;

—M(S): the 'sells' selected by the model;

—U(B): the best performing 20% of shares in the universe from the selection date; and

—U(S): the worst performing 20% of shares in the universe from the selection date.

9.5.3 The performance index I for the model is calculated weekly as:

$$I = \frac{\mathrm{M}(B) - \mathrm{M}(S)}{2}$$

and the '100% hindsight' performance P for the universe is calculated weekly as:

$$P = \frac{\mathrm{U}(B) - \mathrm{U}(S)}{2}.$$

Finally, the 'value added' ratio V is calculated as:

$$V = \frac{I}{P}.$$

9.5.4 The crucial indicator is I, and, in the case of a successful model, it will increase, possibly somewhat erratically, to a value well in excess of that required to cover the level of switching costs appropriate to an institutional investor. Since the value of I will depend on the aggregate price volatility of the universe as well as on the combined predictive power of the model and its users, V is a useful diagnostic test statistic, particularly when different models have been run using different universes of shares.

9.5.5 The results of my equity model, as described in the previous section, give typical values for I of around 0.03 after three months and around 0.045 after six months, which I interpret as a comfortable pass on my 'Turing test'. The challenge that I leave for financial economists is to either confirm or reject my refutation of the Efficient Market Hypothesis, by carrying out their own controlled tests of equity models which are documented in the actuarial literature.

9.6 *Roemer and the Speed of Light*

9.6.1 In 1676 the Danish astronomer Roemer had been making measurements of the longitudes of various places on the earth. Aware of the limitations of seventeenth century clocks, he wanted some absolute measure of time, and thought that the four largest moons of Jupiter could be used as a celestial clock, on the basis that the times between successive eclipses as they moved into the sun's shadow of Jupiter would be constant.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

9.6.2 Despite the inaccuracies of his mechanical clocks, Roemer found that the periods between eclipses changed slowly, but systematically, over time. Using data for Io, the closest of these four moons to Jupiter and the one with the highest frequency of eclipses as a result of its periodic time being shortest, he found that the difference between the longest and shortest eclipse periods was $16\frac{1}{2}$ minutes.

9.6.3 Having encountered this totally unexpected behaviour, Roemer set about trying to find some explanation. On considering the relative positions of the earth and Jupiter as they moved with differing periodic times in very nearly circular orbits around the sun, he realised that his mistake had been to assume that light was an instantaneous effect throughout the solar system. The apparent discrepancy of $16\frac{1}{2}$ minutes clearly related to the time that it took for light to travel the additional distance of the earth from Jupiter as between their closest and most distant relative positions, as shown schematically in Figure 9.1.

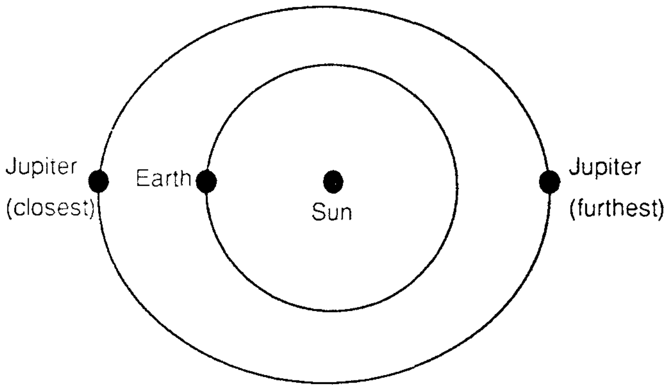


Figure 9.1.

9.6.4 The path, P_t that light has to travel at time t between Io and the earth can be represented in terms of the following simple mathematical model: the

$$P_t = C + D \sin(\mathrm{J}(t)) + E$$

where:

C is the mean distance from Jupiter to the sun;
D is the mean distance from the earth to the sun;
$\mathrm{J}(t)$ is an almost linear function of t involving the relative periodic times of Jupiter and the earth; and
E is a small error term for other effects such as the orbits of Jupiter and the earth not being perfectly circular.

9.6.5 The variation of P_t over time will approximate to simple harmonic

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

motion, with the maximum and minimum values differing by $2D$. Using clocks that were very poor by today's standards, and a somewhat inaccurate seventeenth century value for the distance from the earth to the sun, Roemer obtained a value of 225,000 km per second for the speed of light, which, in the circumstances, compares quite well with the modern value of 299,792 km per second. A detailed description of Roemer's pioneering work on the speed of light can be found in Ronan (1981).

9.7 *Simple Harmonic Motion and Capital Market Prices*

9.7.1 Now consider the simplifying assumption within finance theory as to the essentially instantaneous manner in which capital market prices react, leading to an equilibrium position where the risk-adjusted returns on all shares are equal and where all price changes are the result of new information.

9.7.2 For investment trusts, individual gilts and individual equities, the mean absolute deviation charts described earlier exhibit behaviour that is totally inconsistent with this simplifying assumption of essentially instantaneous adjustment to new information. Not only are the upper and lower control limits (except relatively recently in the case of gilts) sufficiently far apart to allow profitable switching opportunities after transaction costs, but strongly cyclical patterns also exist, whereas the arrival of new information should be purely random over time.

9.7.3 If prices moved in a 'random walk' fashion between 'reflecting barriers', as suggested by Cootner (1962), prices would, on average, be half-way between the central value and one of the reflecting barriers (which would also be the control limits). The mean absolute deviation would then be half the distance between the central value and each control limit, and the constant multiplier would be, accordingly, 2. If there was a strong tendency for prices to gravitate towards the central value, the mean absolute deviation would be smaller and the multiplier would be greater than 2.

9.7.4 In practice, for investment trusts, gilts and equities, not only is the multiplier less than 2, but the same value of 1.6 also works well for all three of these diverse types of security. Furthermore, it is easy to show that the multiplier for a residual exhibiting simple harmonic motion (i.e. a pure sine wave) is $\pi/2$ or about 1.57, which seems too close to the empirical value of 1.6 to have arisen purely by chance.

9.7.5 By analogy with Roemer's elementary model for the solar system to explain an unexpected variation that approximated to simple harmonic motion, we wish to find an elementary model for investor behaviour which will explain variations in capital market prices that can be seen through the 'actuarial telescope' of mean absolute deviation analysis to exhibit properties consistent, in important respects, with simple harmonic motion.

9.8 *The Dynamic Equilibrium Model*

9.8.1 Suppose that we regard type-B investors as 'trend-chasers' who, in the

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

absence of intervention by other types of investor, would cause the rate of change of a share price over time to remain constant over the medium term. Suppose, also, that we regard type-A, type-O and type-R investors as 'centralisers', whose aggregate scale of intervention (i.e. buying when the price is below the central price and selling when above) is proportional to the deviation of the price from its central value. Then, by analogy with the pendulum, this interactive system will generate simple harmonic motion of share prices around their central values of precisely the general type detected by mean absolute deviation analysis.

9.8.2 To determine the dynamics of this share price motion over time, we can use the further analogies of two 'Laws of Motion' that Newton applied so successfully in the physical sciences:

First Law A body at rest or in uniform motion in a straight line will continue in that state until acted upon by an external force.

Second Law Force equals mass times acceleration.

9.8.3 Let the proportion of type-B investors be b, so that the aggregate proportion of type-A, type-O and type-R investors together is $1-b$. Then the 'mass' or inertia of share prices resulting from the actions of type-B investors will be proportional to b, while the centralising force resulting from the combined actions of type-A, type-O and type-R investors will be proportional to $1-b$. The centralising force will thus be proportional to $(1-b)(P_t-C)$, where P_t is the share price at time t and C (regarded for the moment as a constant) is the central price. Also, 'mass times acceleration' will be proportioned to $b\ddot{P}_t$, where $\ddot{P}_t$ is the second derivative of P_t with respect to time. Introducing the square of an arbitrary positive constant k, the equation of motion is now:

$$k^2(1-b)(P_t-C) = -\, b\, \ddot{P}_t.$$

9.8.4 We note, also, that the maximum price is $C+D$, where D (also regarded for the moment as a constant) is the difference between the central value and each control limit. The general solution of the above second order differential equation, subject to this boundary condition involving D, can easily be shown, after introducing an error term, to be:

$$P_t = C_t + D_t \sin\left(\sqrt{\frac{1-b}{b}}k + t_0\right) + E_t$$

which is virtually identical to the model that Roemer used in astronomy more than three centuries ago to determine the speed of light. Since this new model expresses share prices in terms of a dynamic equilibrium position that is caused by the interaction of two quite different forces of investor behaviour, an appropriate description for it would appear to be the 'Dynamic Equilibrium Model'.

9.9 *Predictions of the Dynamic Equilibrium Model*

9.9.1 I now describe, in decreasing order of both obviousness and practical relevance, general predictions as to the behaviour we should expect in the real financial world on the basis of my Dynamic Equilibrium Model.

9.9.2 The first prediction is that the expected return from investment in a particular security will vary with the type of investor, and will decrease successively from the type-R investor to type-O, type-A, and finally type-B. This variation with investor competence, while totally foreign to the teachings of financial economics, is consistent with the comment in Clarkson & Plymen (1988) that "the Perfect Market Principle should not apply to the erudite, professional investor, well equipped and well advised, who should be able to find many opportunities for situations with undervalued and overvalued shares, both among the leading issues and particularly within the smaller companies that get less research attention".

9.9.3 The second prediction is that, other than for type-B investors, the expected return increases as the price volatility, as measured by D_t/C_t, increases. This insight, while the exact opposite of the teachings of financial economics that higher variance is equivalent to higher risk and is therefore always 'bad', is consistent with the following comments in Clarkson & Plymen (1988):

> "For the longer term, it is the growth rate that determines the risk, the variance being of negligible importance. A volatile price structure can even be an advantage, if the share can be bought on cheap terms with consequent lower eventual risk."

9.9.4 The third prediction is that, for a reasonably homogeneous sample of investment funds, short-term volatility of return, as measured by standard deviation of return, will vary in a U-shaped manner with relative performance, since funds that outperform or underperform significantly will have far more upside and downside deviations, respectively, than funds performing roughly in line with the sample average. This is illustrated below for the 100 North American unit trusts with a five-years' performance record to 1 July 1995. The performance rankings over five years and the standard deviations of return over two years are from the August 1995 issue of *Money Management.*

Performance ranking	Average standard deviation (% p.a.)
1 — 10	3.68
11 — 20	3.43
21 — 30	3.45
31 — 40	3.18
41 — 50	2.98
51 — 60	2.69
61 — 70	2.86
71 — 80	2.92
81 — 90	3.00
91 — 100	3.33

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

This pattern is inconsistent with one of the main teachings of financial economics, namely that investments with the lowest expected returns will also have the lowest risk, as measured by standard deviation of return.

9.9.5 My fourth prediction is that the price variability of equities will vary systematically with the difficulty of predicting a company's future earnings. Thus utilities, where the overall profitability is generally controlled by government regulation, will exhibit relatively low variability, whereas oil exploration stocks and high technology stocks will have highly unpredictable profit streams, and hence very high price variability. This behaviour is demonstrated very clearly in Osborne (1959), where oil and mineral exploration stocks traded in Toronto had, over all time periods investigated, vastly higher price variability than utility stocks traded in New York.

9.10 *The Actuarial Telescope*

I see strong parallels between the use of the telescope in astronomy some four centuries ago and mean absolute deviation analysis as an actuarial investment tool. The greatly enhanced understanding of planetary behaviour, made possible by the telescope, proved, beyond all reasonable doubt, the untenability of the Ptolemaic system of astronomy. Similarly, the greatly enhanced understanding of capital market behaviour, made possible by mean absolute deviation analysis, should, I believe, be accepted as convincing evidence of the untenability of current theories of financial economics in the areas of equilibrium and rapid assimilation of new information.

9.11 *The Sharpe Diagonal Model*

The Sharpe Diagonal Model, which is essentially static in nature, can be expressed as:

$$R = A + B\,I + C$$

where:

- R is the return on a security over an arbitrary period;
- A is the 'excess return' or 'alpha';
- B is the 'systematic risk' or 'beta';
- I is the market return; and
- C is a random error term.

The model depends very much on a simplistic statistical interpretation of the past, rather than on comparisons of current relative ratings against future prospects, and it cannot take into account the observed cyclicality of capital market behaviour. In particular, attempts to apply the model in practice will tend to reinforce 'unintelligent' trend-chasing behaviour by suggesting, regardless of any long-term appraisals, the purchase of 'high beta' securities in rising markets and their sale in falling markets.

9.12 *Look after the Expected Return and the Risk will look after Itself*

9.12.1 The main conclusion of Clarkson & Plymen (1988) is that "if advanced analytic techniques can assess the likely returns on shares more accurately than most market participants, then the probability distribution of the return on a portfolio 'moves to the right' and improves the expected return while simultaneously reducing the risk".

9.12.2 In this context, risk is used in the obvious, common sense interpretation of performance falling below some benchmark value such as a market index, the average or median performance of competitors, or a specific benchmark set by a client. Whatever the benchmark, it is intuitively obvious that the probability profile of returns assuming the use of advanced techniques involves lower risk than the other profile.

9.13 *A Fundamental Axiom of Risk*

These considerations suggest in 'pattern recognition' terms the following fundamental axiom of risk:

— If, for two probability profiles of return, one involves a lower probability than the other for all returns below the risk threshold, then the former has the lower risk.

9.14 *A Tentative New Measure of Risk*

9.14.1 The above axiom, though perhaps a small step forward, is much too restrictive to be of any practical use. Accordingly, as a tentative first step towards replacing variance as the universal measure of risk, we suggest, for two probability profiles A and B, a shortfall measure of risk relative to profile B, namely the probability-weighted square of deviations below the expected return on profile B.

9.14.2 This can immediately be generalised for any probability density function of return $p(r)$ and risk threshold L to:

$$R = \int_{-\infty}^{L} (L - r)^2 p(r) dr$$

which is the generalisation of semi-variance that results from the use of a variable risk threshold L in place of the expected return.

9.15 *Practical Difficulties*

We stress that these tentative suggestions are theoretical, rather than practical, in nature, since the underlying probability density functions are so complex, and depend on so many economic variables, that further analysis along these lines is pointless in most practical applications.

9.16 *Conclusions*

The more we study the methodologies of MPT, the stronger is our belief that

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

'traditional' methods of security analysis and portfolio construction are of far greater practical relevance. The Dynamic Equilibrium Model is, accordingly, put forward, not as a practical methodology for investment professionals, but as a less unrealistic conceptual framework against which financial economists can test the validity, or otherwise, of what we perceive to be the very unsound theoretical foundations of Modern Portfolio Theory and similar quantitative approaches.

10. THE MEASUREMENT OF INVESTMENT RISK

10.1 *Introduction*

10.1.1 It proved far more difficult than I had expected to find a 'first principles' approach to investment risk that would result in a formula of the type:

$$R = \int_{-\infty}^{L} (L - r)^2 \, p(r) dr$$

and my paper (Clarkson, 1989) entitled 'The Measurement of Investment Risk', as presented to the Faculty of Actuaries in March 1989, ran to more than fifty pages, as against my original estimate of ten. I had, in fact, to go right back to first principles by using my personal experience of potentially dangerous sports, such as rapid river canoeing and ski-mountaineering, as pointers to a general approach to risk. Only after I had constructed a framework for physical risk in sports did I attempt to construct a corresponding framework for financial risk.

10.1.2 Although I had extreme difficulty in selecting probabilities for my numerical examples relating to physical risk in sports, I made no explicit reference to this. Nor did I mention that, despite my training, both as a mathematician and as an actuary, most of the critical decisions that I had to make in high-risk sports situations, including some that were in the 'life-or-death' category, had virtually nothing to do with the conscious and algorithmic mental processing of 'known probabilities' as derived from past experience. My implicit assumption of 'known probabilities' in sports is very much like what Tolstoy describes as the fiction of 'vital force'; it might, as Adam Smith pointed out, soothe our anxiety to some extent, but it will fail to get to the heart of the problem. As a result, we risk misleading ourselves as to whether we have, indeed, found a useful 'scientific' explanation for a perplexing real world phenomenon. Had I read Penrose's attack on 'purely algorithmic thought' twelve months before, rather than, as was the case, twelve months after, writing the paper, I would have stressed that my experience fully supported his controversial conclusions as to the functioning of the human mind, and, accordingly, that any approach to risk based on 'known probabilities' would be seriously incomplete in terms of explaining human behaviour in the real world.

10.2 *Likelihood and Severity*

10.2.1 In assessing physical risk in sports, it is clear to me that 'likelihood' and 'severity' have to be assessed separately over possible scenarios that could constitute an 'adverse occurrence'. In the case of hang-gliding, an adverse occurrence involves a stall or an equipment failure when high above the ground, and is likely to result in death; in wind-surfing, it involves no more than losing control, falling into the water, and having to clamber back onto the board. The probability of an 'accident' cannot, therefore, be used in any meaningful way as a measure of risk.

10.2.2 Assessments of likelihood and severity on a very general basis lead to the following table for different sports:

	Adverse occurrence		Likelihood of adverse occurrence resulting in:	
Sport	Nature	Likelihood	Serious injury	Death
Hang-gliding	Plunge to ground	very low	very high	very high
Rapid river canoeing	Capsize in rapids	moderate	low	low
Ski-ing	Bad fall	moderate	moderate	very low
Ski-mountaineering	Avalanche	low	high	high
	Bad fall	moderate	moderate	low
Wind-surfing	Falling off	very high	negligible	negligible

10.2.3 An obvious next step is to regard the 'contribution to risk' for a possible outcome (e.g. serious injury) as the product of the probability of that outcome arising and a weighting function that reflects, in psychological terms, the severity of that outcome. The final measure of risk is then the sum over all possible outcomes of these contributions to risk. On this basis, and since the perceived degree of risk is highly dependent on the most serious possible outcomes, the information in the above table suggests that wind-surfing will have the lowest risk, hang-gliding and ski-mountaineering will be the two with the highest risk, and ski-ing and rapid river canoeing will be somewhere in between.

10.3 *Introducing a Risk Threshold*

10.3.1 If we take a group of people who all agree on this overall ranking, some will be more risk-averse than others, as evidenced by where their 'cut-off point' of riskiness corresponds to the perceived risks of the various sports. Most able-bodied individuals would, for instance, be prepared to accept the risks involved in ski-ing, but only the most adventurous would be prepared to participate in hang-gliding.

10.3.2 There is clearly a risk (however small) of serious injury or death in all day-to-day activities. Most people are prepared to fly on normal passenger aircraft for holiday and business purposes, and virtually everyone is prepared to travel by car, although it is evident that there is a non-zero risk of death in each case. These risks are perceived as being below the 'cut-off point' of riskiness that they are prepared to accept.

10.4 *Principles, a Simple Model, and Axioms*

10.4.1 The approach to risk underlying this sports example suggests the following principles:

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

(1) The perceived degree of risk relates both to the severity of the consequences of an adverse occurrence and also to the probability of these consequences arising.
(2) The perceived degree of risk is highly dependent on the probabilities of the most serious possible consequences.
(3) In assessing whether a particular activity is acceptable in terms of the attendant risk, an individual will compare the perceived risk with some threshold of riskiness based on his or her personal preferences and experience.
(4) Individuals may differ in the degree of risk they are prepared to accept.

10.4.2 Using these principles, we can translate the sports example into a simple mathematical model which will help an individual decide which sports are admissible, or inadmissible, in risk terms, in the light of his or her personal threshold of risk. The first step is to classify possible outcomes in terms of a numerical scale, and the scale I use is:

0	No injury
1	Minor injury
2	Moderate injury
3	Severe injury
4	Very severe injury
5	Permanent incapacity
6 and higher	Death

10.4.3 A very convenient class of probability distributions to link in with this scale is the Poisson distribution (a limiting case of the binomial distribution), where the parameter λ (which represents both the expected value and the variance) increases with the likely severity of the consequences. A method of weighting for the different degrees of severity is also required. A plausible choice here is 1 for death, 0.1 for permanent incapacity, 0.01 for very serious injury, etc., so that each next lower degree of severity decreases the contribution to risk by a factor of 10. The resulting risk value can then be interpreted as the 'equivalent probability of death', where the underlying psychological impact is based on equivalences such as 'death is equivalent to 100 very severe injuries'.

10.4.4 For ski-mountaineering, using $\lambda = 3$ for the avalanche risk and $\lambda = 0.2$ for the bad fall risk, the risk values in 'parts per million' per occurrence are calculated as follows:

Severity	Probability	Weight	Contribution to risk	Probability	Weight	Contribution to risk
0	0.0498	0	0	0.8187	0	0.00
1	0.1494	10	1	0.0164	10	1.64
2	0.2240	100	22	0.0164	100	1.64
3	0.2240	1,000	224	0.00109	1,000	1.09
4	0.1680	10,000	1,680	0.000055	10,000	0.55
5	0.10082	100,000	10,082	0.0000022	100,000	0.22
6 and higher	0.083918	1,000,000	83,918	0.0000007	1,000,000	0.07
			95,927			5.21

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

10.4.5 Assuming one chance in a thousand of an avalanche each day, and two bad falls per day, the final value of risk (in parts per million) is:

$$95{,}927 \times 0.001 + 2 \times 5.21$$
$$= 95.9 + 10.4$$
$$= 106.3.$$

We can then say that the risk, as the equivalent probability of death per day, is about 1 in 10,000, with around 90% of this arising from the avalanche risk.

10.4.6 Using appropriate values for λ and the daily occurrence rates in the other sports gives daily risk levels which confirm the general pattern expected in ¶10.2.3:

Sport	Risk (parts per million per day)
Wind-surfing	0.5
Ski-ing	3.1
Rapid river canoeing	13.8
Ski-mountaineering	106.4
Hang-gliding	142.6

10.4.7 To complete the example, we require some method of calculating a 'cut-off' value of risk. An obvious approach here is to specify that the probability of an accident of a certain severity (or worse) should not exceed some chosen small quantity. A reasonable choice is 'not more than one chance in a thousand per day of serious injury or worse'. To translate this into a value of risk, we require to know the relative probabilities of the various levels of severity, and a satisfactory assumption here is that they follow the same pattern as for a 'medium risk' sport (i.e. rapid river canoeing), which is in accordance with a Poisson distribution with $\lambda = 0.5$. Since the risk for a single adverse occurrence on this basis is 69.0×10^{-6} and the probability of an outcome of severity 3 or higher is 0.01439, the daily level of risk on the specified criterion is:

$$69.0 \times 10^{-6} \div 14.39, \quad \text{i.e. } 4.8 \times 10^{-6}.$$

10.4.8 An individual who specifies this threshold of risk would, therefore, only consider participating in wind-surfing and ski-ing; the other three sports are too risky. Since rapid river canoeing is viewed by most people as a fairly risky sport, this threshold value of daily risk of around 5 parts per million is probably of the right order of magnitude for the vast majority of individuals who do not wish to participate in sports that are perceived to involve any meaningful degree of risk. However, since I enjoy ski-mountaineering, but regard hang-gliding as too risky, my personal threshold of daily risk, on the basis of the risk values that I have obtained, must be around 125×10^{-6}, i.e. half way between the risk values for ski-mountaineering and hang-gliding.

10.4.9 We can now formulate the following general axioms regarding risk:

(1) Risk is a function both of the probabilities of possible adverse outcomes and also of the severity of the consequences of these adverse outcomes.

(2) If the probabilities of the possible adverse outcomes are known, risk can be expressed as a non-negative measure:

$$R = \int W(s)p(s)ds \text{ or } \sum W(s_i)p(s_i)$$

where $p(s)$ (or $p(s_i)$) relates to consequences of severity s (or s_i) and $W(s)$ is a positive function of s which increases with s.

(3) For a given value of the measure of 'enjoyment' E, an individual will prefer the course of action with the lowest risk.

(4) Each individual has a threshold of risk R_0, and will not pursue a course of action which involves a value of risk higher than R_0.

(5) An individual will choose between all possible courses of action by maximising E subject to the risk not exceeding R_0.

(6) Different individuals may differ in their degree of aversion to risk by using different functions $W(s)$ and/or different risk thresholds R_0.

10.5 *A Sports Paradox*

10.5.1 Although this essentially probabilistic approach appears to offer a useful way forward, it is far from being a complete theory of how the human mind first of all measures, and then manages, risk. Suppose that I accept the risk values for each of the sports individually and also my personal daily risk threshold of 125×10^{-6}. I am now confronted with two separate choices:

Choice 1

A a 50:50 chance of either wind-surfing or hang-gliding, the decision to be made by the toss of a coin; or

B ski-mountaineering with certainty;

Choice 2

C wind-surfing with probability 0.99 and hang-gliding with probability 0.01, the decision to be made by drawing one of 100 lottery tickets; or

D rapid river canoeing with certainty.

10.5.2 I have no hesitation in choosing B in preference to A and D in preference to C, and in both cases the 'enjoyment' element does not enter into my thought processes; I regard A and C as too risky for me, even although, on a look-through basis, A involves a lower risk than B, C involves a lower risk than D, and both A and C are well below my personal risk threshold. The paradox is that my thought processes, in terms of assessing the risks of compound events, are inconsistent with the axioms of probability. I shall present my suggested resolution of this paradox after developing my framework for financial risk.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

10.6 *Translation into Financial Risk*

10.6.1 In translating the general axioms of physical risk into provisional axioms that apply to investment, we note the following:

(a) Since investment returns are essentially continuous in nature, an integral measure rather than a summation is appropriate.
(b) If $p(r)$ is the density function of return of the investment being considered, the severity of the adverse consequences increases as r decreases.
(c) There will be a value L or r above which there are no adverse consequences.
(d) The obvious measure of 'enjoyment' is the expected return $E(r)$.

10.6.2 The provisional axioms of investment risk are thus:

(1) Investment risk is a function both of the probability of the return being below a certain threshold and also of the severity of the financial consequences of these values of return.
(2) If the density function of return, $p(r)$, is known, risk can be expressed as a non-negative measure:

$$R = \int_{-\infty}^{L} W(L-r)p(r)dr$$

where $W(s)$ is a positive function of s which increases with s and L and is the value of return above which no adverse consequences arise.
(3) For investments with the same expected return, an investor will prefer the investment with lowest risk.
(4) Each investor has a threshold of risk R_o and will not make an investment which involves a value of risk higher than R_o.
(5) An investor will choose between all possible investments by maximising the expected return subject to R not exceeding R_o.
(6) Different investors may differ in their degree of aversion to risk by using different functions $W(s)$ and/or different thresholds or risk R_o.

10.6.3 Two observations can be made at this stage. First, the risk measure R in (2) is already tantalisingly close to the integral expression at the beginning of this section. Second, (5) is a highly arbitrary, and also very restrictive, interpretation of 'intelligent' behaviour in the face of risk and uncertainty; a more general axiom will clearly have to be found before we can claim to have a satisfactory framework for financial choice under uncertainty.

10.7 *Risk Weighting Function*

10.7.1 Consideration of the general properties of the risk weighting function $W(s)$, using simple numerical examples, suggests that $W(s)$ will be a function of the degree of aversion to increased risk perceived by an individual. It will, therefore, be desirable to use a one-parameter family of functions $W_a(s)$, where a is the parameter describing the degree of aversion to increased risk and

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

$\frac{1}{W}\frac{dW}{ds}$ increases with a.

10.7.2 The two most obvious simple classes of function satisfying these requirements are:

$$W_a(s) = s^a \text{ for } a > 1$$

and

$$W_a(s) = e^{as} \text{ for } a > 0$$

and either class of function could offer a useful theoretical framework.

10.8 *Commutation Functions*

The former class of function has highly desirable properties that simplify the computational aspects very significantly. In particular, for a particular 'shape' of probability distribution of return, it is possible to define a 'risk commutation function' R(0,1) for the distribution with mean 0 and variance 1. Risk values for other values of mean and variance can then be calculated from appropriate values of R(0,1). For these and similar practical considerations, it seems appropriate to base the general framework for financial risk on the former, rather than the latter, class of risk weighting function.

10.9 *Axioms of Financial Risk*

10.9.1 The general framework for financial risk can now be expressed in terms of eight axioms:

Axiom 1 Investment risk is a function both of the probability of the return being below a certain threshold and also of the severity of the financial consequences arising from these values of return.

Axiom 2 If the density function of return $p(r)$ is known, investment risk can be expressed as a non-negative measure:

$$R = \int_{-\infty}^{L} W\left(\frac{L-r}{k}\right)p(r)dr$$

where:

L is the value of return above which no adverse consequences arise;

$W(s)$ is defined for $s > 0$;

$W(s) > 0$;

$\frac{dW}{ds} > 0$;

$\frac{d^2W}{ds^2} > 0$; and

k is a positive scale parameter.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

Axiom 3 For investments with the same expected return, an investor will prefer the investment with the lowest risk.

Axiom 4 Each investor has a threshold of risk R_0, and will not make an investment which involves a value of risk higher than R_0.

Axiom 5 An investor will choose between all possible investments by maximising the expected return subject to the risk not exceeding R_0.

Axiom 6 Different investors may differ in their degree of aversion to risk by using different functions W(s) and/or different thresholds of risk R_0.

Axiom 7 For most practical and theoretical work there is very little loss of generality in assuming:

$$\mathrm{W}\left(\frac{s}{k}\right)=\left(\frac{s}{k}\right)^{a} \quad \text{where} \quad a>1.$$

Axiom 8 For most practical work it will be possible to calculate an explicit value for R_0 by specifying that the probability of financial consequences of a certain degree of severity, or worse, is not to exceed some chosen value, and that the relative frequencies of the various adverse consequences follow a given probability distribution.

10.9.2 Axiom 5 is still very tentative in nature; a suitably generalised axiom will be suggested later.

10.10 *Semi-Variance as a Special Case*

The somewhat restrictive special case of my downside risk measure with the mean of the distribution as the risk threshold and a value of 2 for the risk weighting parameter gives semi-variance, the statistical risk measure which Markowitz regarded as more appealing than variance on intuitive, rather than computational, grounds. However, the use of semi-variance, which cannot take explicit account of the wealth of the individual, is a retrograde step in the context of Bernoulli's initial insight, more than a quarter of a millenium ago, that current or expected wealth should be a crucial determinant of financial behaviour under conditions of risk and uncertainty.

10.11 *Comparisons with Markowitz (1959)*

The general portfolio selection approach, set out in Markowitz (1959), can be summarised in terms of a set of axioms very similar to those in Section 10.9, the only significant difference being that there is no equivalent of my Axiom 5 regarding a maximum level of risk. The theoretical reasons for this and the crucial practical implications are discussed in later sections.

10.12 *Unification*

Parallels with physical risk in sports, where the weighting function can be assumed to remain constant over all outcomes which result in death, suggest that,

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

in the case of financial risk, the weighting function should remain constant over all outcomes which result in appropriate instances of 'financial ruin', such as statutory insolvency or bankruptcy. This introduction of a limiting value makes it possible to unify, in theoretical terms, the apparently diverse approaches to risk used in life assurance, general insurance, moderate risk areas of financial activity such as portfolio selection, and high-risk areas of financial activity such as trading in derivative securities.

10.13 *Practical Considerations*

10.13.1 The unification achieved above is of only theoretical interest, since it ignores the very considerable practical difficulties involved in obtaining appropriate probability density functions. Even in the simple sports example described earlier, I had extreme difficulty in choosing, through the 'hidden variables' of the daily average and the Poisson parameter, reasonably suitable descriptions of the underlying probabilities.

10.13.2 In the case of ski-ing, I first of all assumed two falls per day as a 'beginner' average and then calibrated the Poisson parameter against personal observations over a number of years, which suggested that, on an aeroplane carrying skiers back to the U.K. after ski holidays in Europe, around 60% of the time there were either one or two passengers with a leg or an arm in plaster. I calculated my 'exposed to risk' on the basis that around two thirds of the 120 or so skiers on board would have been abroad for one week (ski-ing for 6 days) with the remainder having been abroad for two weeks (ski-ing for 12 days), thereby giving an 'exposed to risk' of around 960 ski days.

10.13.3 It is interesting to note that an investigation into fatality rates in sports as part of a wider study of risk by the Royal Society (1993) identifies the same major difficulties that I encountered:

(1) differences in the training, experience and competence of the participants; and
(2) differences in the severity of the challenge undertaken.

The Royal Society results were expressed as deaths per 10^6 participant-hours, whereas my risk values are daily participation rates covering death and other outcomes, but with a high weighting towards death. Three sports — ski-ing, canoeing and hang-gliding — are covered in both studies. Using 5 hours, $2\frac{1}{2}$ hours and 15 minutes, respectively, as the average daily participation in these three sports, to convert the Royal Society results to a daily participation basis, gives the following table:

Sport	Clarkson (1989) Daily risk x 10^6	Royal Society (1993) Deaths per day x 10^6	
Canoeing	3.1	2.5	U.K., 1960-62
Ski-ing	13.8	3.5	U.S.A., 1967-68
		6.5	France, 1974-76
Hang-gliding	142.6	100-325	U.S.A., 1978
		375	U.K., 1977-79

10.13.4 I see the surprisingly close agreement between my 'everyman' values and the Royal Society results as very strong evidence that the human mind, in the absence of explicit numerical data, can assess the general level of physical risk fairly accurately through the application of 'pattern recognition thought' to personal experience. In practice, however, the human mind will assess each risk in terms of the strength of a psychological stimulus, so that, in our conscious thinking, we are not aware of a quantitative value of risk on any particular numerical scale.

10.13.5 The widely differing risk values for ski-ing as between the U.S.A. and France are also, I believe, very significant. While a statistician with no practical experience of ski-ing might conclude that the training, experience and skills of skiers in the U.S.A. are much superior to those of skiers in France, I would suggest a quite different causal mechanism, namely the general geography of the mountains on which ski-ing takes place, and, in particular, the level of difficulty involved in ski-ing back to town at the end of the day or when weather conditions deteriorate markedly. At major U.S. ski resorts, there is always an 'easiest' run back to safety. Moreover, while some very difficult runs exist, they have to be deliberately sought out rather than being major access routes from point A to point B. In the Alps, however, the steepness of many slopes back to town and pistes linking one area to another on the same circuit, lead to a far larger proportion of the ski-ing population being exposed to the hazards of very difficult runs.

10.14 *Psychological Considerations*

10.14.1 From my experience of sports, I am aware of a serious impediment to rational thought in high-risk situations that does not seem to be recognised in the economics literature. When the human mind perceives serious risk to life or limb, either in the immediate future or in the recent past, the central nervous system activates the release of adrenalin to put the body into 'over-drive' for the 'fight or flight' situations that posed a serious threat to the survival of our distant ancestors. I am fairly certain that high levels of financial risk cause the same, but in this case totally inappropriate, physiological reactions.

10.14.2 I can now present my suggested resolution of the sports paradox described in Section 10.5, namely that my choices over probabilistic mixtures of sports are apparently inconsistent with the axioms of probability. The crucial new principle that I incorporate in my framework for risk is:

— Do not pursue a course of action which involves any meaningful likelihood of participating in an activity which you perceive as having a risk greater than your personal threshold.

This principle seems essential, in view of the inability of the human mind to function normally when the perceived probability of death, either in the immediate future or the very recent past, exceeds some threshold value that could be as low as 0.001. Complete paralysis of the faculty for rational thought is likely when this probability is of the order of 0.1 or higher. The resolution of the apparent paradox involves recognising this discontinuity in the human mind's capacity to deal with increasing levels of risk. Formal systems of utility theory, on the other hand, incorporate apparently innocuous continuity axioms which imply that the psychological factors that I have described play no part in real world human behaviour.

10.15 *Implications for Rational Behaviour*

10.15.1 I shall develop my arguments in terms of ski-ing, which will be familiar to most readers. For a given level of 'achievement' in terms of difficulty of terrain successfully negotiated, the above psychological considerations suggest that the overall 'enjoyment value' will decrease as perceived risk increases. Also, the above practical considerations suggest that, for the same level of difficulty of terrain, the perceived risk will decrease as the training, experience and skills of the skier improve.

10.15.2 If the 'enjoyment value' of ski-ing a particular run depended only on the achievement level as measured by technical difficulty, the situation would be that implied by Axiom 5, namely vertical indifference lines in the risk-achievement diagram below the level of maximum risk. However, the psychological considerations discussed above suggest that 'enjoyment value' will decrease markedly as the level of risk approaches the maximum, and, accordingly, that a more plausible pattern is indifference curves which:

(1) are asymptotic to vertical straight lines for small values of risk;
(2) decrease in gradient as risk increases; and
(3) are asymptotic to the horizontal straight line represented by the maximum level of risk.

10.15.3 We can now replace Axiom 5 by the following more general axiom:

Axiom 5G An investor will select between investments which have a lower value of risk than his risk threshold, on the basis of risk-return indifference curves which exhibit the following properties:

(1) for small values of risk, the curves are asymptotic to vertical straight lines;
(2) the curves decrease in gradient as risk increases; and
(3) the curves are asymptotic to the horizontal straight line which represents his risk threshold.

10.16 *Conclusions*

My analogies with physical risk in sports lead to a new downside approach to risk, which incorporates an explicit maximum level of acceptable risk. This approach could offer, in the area of financial risk, a much more satisfactory conceptual framework than the 'risk equals variability of return' paradigm of financial economics.

11. The Assessment of Financial Risk

11.1 *Introduction*

Clarkson (1990) was written to investigate possible applications of the downside approach to risk described in the previous section, with particular emphasis on the St Petersburg Paradox, the Black-Scholes option pricing formula, and risk-return comparisons of different investment distributions.

11.2 *Indifference Curves*

The first step is to specify indifference curves in the risk-return diagram that are consistent with Axiom 5G. I use a one-parameter family of hyperbolae with a value of 1.5 for the parameter b which determines the convexity. For the risk weighting parameter I use values of 1.5, 2, 2.5 and 3 for low, moderate, high and very high degrees of aversion to increasing risk, respectively.

11.3 *Bernoulli and the St Petersburg Paradox*

11.3.1 Suppose that A tosses a coin until heads appear. If it shows heads on the first throw, he receives 1 from B, if it does not show heads until the second throw he receives 2 from B, if it does not show heads until the third throw he receives 4 from B, and so on, with A receiving 2^{N-1} from B if it first shows heads on the Nth throw. How much should A pay B to enter this game?

11.3.2 On the basis of classical probability theory alone, the expected value of the payment to A is infinite, whereas common sense suggests that no reasonable person would pay B more than a modest amount. The paradox is then that probability theory is, apparently, directly opposed to the dictates of common sense.

11.3.3 The suggested resolution of this paradox by Daniel Bernoulli (1738) using 'moral expectation' (or what we would nowadays call a logarithmic utility function) marked the beginning of utility theory. It is, therefore, appropriate to compare Bernoulli's approach with my more general framework for risk.

11.3.4 Bernoulli begins by observing, as mentioned in ¶6.11.3, that a poor fellow obtaining a lottery ticket that will yield, with equal probability, nothing or 20,000 ducats would be stupid not to sell it for 9,000 ducats, while a rich man would be ill-advised not to buy it for 9,000 ducats. Bernoulli concludes from this that individuals with differing circumstances should evaluate the gamble in different ways. He then makes very sweeping simplifying assumptions, by considering only increases in wealth that proceed continuously by infinitesimal increments, and stating that, if one individual "has a fortune worth 100,000 ducats and another a fortune worth the same number of semi-ducats, it is quite clear that to the former a ducat has exactly the same significance as a semi-ducat to the latter, and that, therefore, the gain of one ducat will have to the former no higher value than the gain of a semi-ducat to the latter."

11.3.5 Having translated these relationships for increasing wealth into a logarithmic utility curve above the point of current wealth, he then makes the critical assumption that the same logarithmic curve can be extended to describe preferences below the point of current wealth "since in a fair game the disutility to be suffered by losing must be equal to the utility to be derived by winning." This assumption is inconsistent with my postulate of a limiting value of risk above which an individual will not willingly go.

11.3.6 Using my standard risk parameters of a = 2 and b = 1.5 and a plausible formulation for the decreasing marginal utility of wealth effect, the values that I obtain, with and without allowing for risk, together with the values obtained by Bernoulli, are:

A's wealth as multiple of unit value	Payment to B as multiple of unit value		
	Ignoring risk	Allowing for risk	Bernoulli
10	4.00	2.98	3.04
100	5.63	5.48	4.39
1,000	7.24	7.24	5.97

It must be stressed that the values ignoring risk are strongly dependent on the formulation chosen for the decreasing marginal utility of wealth effect. The values allowing for risk are dependent, also, on the assumptions regarding A's attitude to risk; someone with a low degree of aversion to risk might regard it as rational to pay an amount closer to that calculated above ignoring risk, while a very cautious individual might pay far less than the amount calculated above allowing for risk.

11.3.7 My overall conclusions are that Bernoulli's suggested solution is not only too inflexible, but is also dominated by the decreasing marginal utility of wealth effect for possible outcomes above the value of current wealth rather than being a reflection of reasonable risk-reward preferences. These conclusions are mirrored by Keynes (1921), who, after discussing the very significant elements of psychological doubt, observes that "the theoretical dispersal of what element of paradox remains must be brought about, I think, by a development of the theory of risk".

11.3.8 Bernoulli's suggested numerical solution is also far beyond the grasp of the normal human mind. Although I have not seen any comment on this crucial practical aspect in the theory of finance literature, the computational problems in obtaining Bernoulli's values of 3.04, 4.39 and 5.97 are formidable.

11.4 *The Black-Scholes Option Pricing Formula*

11.4.1 Since I had always regarded the simplifying assumptions used to justify the Black-Scholes option pricing formula as unrealistic in the extreme, I was curious to see how the general pattern of option prices derived from my risk framework would compare with the Black-Scholes value. To my great surprise, the two sets of values, for the same underlying probability distributions of return, were, for all practical purposes, identical. In Clarkson (1990) I show the following values for a one-year call option with different values of K, the ratio of the exercise price to the current price:

K	Clarkson	Black-Scholes
0.8	0.281	0.281
0.9	0.212	0.212
1.0	0.157	0.157
1.1	0.113	0.111
1.2	0.080	0.078

11.4.2 My tentative conclusion was that the Black-Scholes formula gives the correct answer, but for quite the wrong reasons. This conjecture is discussed in Section 15.

11.5 *Textbook Inconsistencies*

11.5.1 Example 1.1 of the actuarial risk theory textbook Bowers *et al.* (1986) involves a comparison of the risk/return characteristics of two different random economic prospects, a normal distribution with mean 5 and variance 2, and a normal distribution with mean 6 and variance 2.5. Since the latter 'lies to the right' of the former, it will have the lower value of downside risk on my approach for any reasonable values of L and a, the risk threshold and risk weighting parameters respectively. Using 7-point binomial distributions as fairly accurate statistical approximations to these two normal distributions, and taking 6 as the risk threshold, we can very easily calculate the risk values for different values of the risk weighting parameter as follows:

	Risk values		Ratio
a	N(5,2)	N(6,2.5)	
1.5	1.766	0.852	2.07
2.0	2.739	1.250	2.19
2.5	4.436	1.909	2.32
3.0	7.445	3.026	2.45

11.5.2 Similar ratios are obtained using other values of the risk threshold. I therefore conclude that, on any common sense interpretation, the distribution N(6,2.5) has a risk value somewhat less than half that of the distribution N(5,2). Since N(6,2.5) also has a higher mean than N(5,2), I conclude that the distribution N(6,2.5) is preferable to the distribution N(5,2).

11.5.3 The 'solution' to this example uses an exponential utility function to 'prove' that the distribution N(5,2) has a higher value of expected utility than the distribution N(6,2.5), and hence is preferable, the heuristic justification being that the more diffuse nature of the distribution N(6,2.5) is a highly adverse factor. I regard this as a classic example of the dangerously unsound conclusions that can result from a combination of utility theory and the 'risk equals variability of return' paradigm of financial economics.

11.6 *Observations by Allais and Sharpe*

11.6.1 If the theoretical foundations of utility theory are, indeed, unsound, one would expect eminent economists to have come to a similar conclusion and to have recorded their views in the finance theory literature. I refer here to some observations along precisely these lines by two winners of the Nobel Prize for Economics.

11.6.2 Perhaps the most detailed refutation of the so-called theoretical foundations of utility theory is given by Allais (1953). Although the main paper is in French, a few passages from the English summary are sufficient to demonstrate the severity of his criticisms.

> "According to the American school, a rational man must conform to the principle of Bernoulli. In our view, this is a mistake which in fact is tantamount to neglecting the fourth specific element in the psychology of risk."

> "Whatever their attraction might be, none of the fundamental postulates leading to the Bernoulli principle as formulated by the American school can withstand analysis. All are based on false evidence."

> "For the rational man, there does not exist in general an indicator $B(x)$ such that the optimum situation could be defined by maximising the expected value of $B(x_i)$."

11.6.3 Sharpe (1970) investigates utility theory as a plausible framework for the implementation of the mean-variance approach to investment selection, and observes that, of the various possible utility curves that have been proposed, "Only one is completely consistent with choices based solely on expected return and standard deviation of return: the assumption that utility is a quadratic function of wealth." However, on investigating the implications of using a quadratic utility curve, he discovers some serious inconsistencies and draws the following conclusions:

> "In some instances, investors will be concerned with more than the expected return and standard deviation of return. In such cases a quadratic utility curve will imperfectly approximate an investor's actual utility curve. If portfolios with radically different prospects are considered by an investor, too much reality may be omitted if his decision is assumed to depend only on expected return and standard deviation of return."

11.6.4 Despite these obvious inconsistencies, Sharpe suggests that the use of a utility curve may still be justified if it is assumed that investors choose amongst portfolios of roughly similar risk:

> "But few investors seriously consider portfolios with radically different prospects. Most people prefer well-diversified portfolios."

I am in general agreement with this pragmatic viewpoint in the very specific context to which Sharpe refers, but in many suggested applications of utility theory (e.g. whether, or not, to insure against a serious financial risk) the risk levels of the scenarios being compared differ enormously.

11.7 *Conclusions*

The downside approach to risk, built up from first principles in Section 10, appears to be more satisfactory, in many regards, than utility theory, and, in particular, it provides a plausible causal mechanism for the Black-Scholes option pricing formula, which is generally regarded as the most successful application of mathematics within modern finance theory.

12. A Non-Linear Stochastic Model for Inflation

12.1 *Introduction*

12.1.1 In the discussion at the Faculty on the comprehensive stochastic investment model described in Wilkie (1986), I was very critical of the use of a linear ARIMA time series for the crucial inflation series which drives the whole model. Having discovered, many years earlier, that the Box-Jenkins ARIMA methodology for projecting a series one step ahead could not even be used to predict the general behaviour of the relative price residuals of individual gilt-edged stocks, I observed that any linear model would be 'far too tame' to encapsulate the behaviour of inflation, which is clearly a much more complex and much more erratic time series, and I suggested the outline of an alternative inflation model that included an intermittent upward-only shock term.

12.1.2 After studying stochastic models in more detail while a member of the Wilkie Review Group in 1989 and 1990, I described, in Clarkson (1991), the construction of a non-linear stochastic model for inflation which incorporated three quite distinct non-linear components and showed that this model gave a better representation of the post-war U.K. inflation experience than the ARIMA inflation model described in Wilkie (1986).

12.2 *Uncertainty of the First Kind*

12.2.1 My 'far too tame' comment relates partly to what I describe as 'uncertainty of the first kind'. In principle, this can be dealt with, as in the Robinson Crusoe example, by running the stochastic model separately for 'high' and 'low' scenarios in addition to the 'central' scenario.

12.2.2 It is important to note that this 'central' scenario is different in concept from what Wilkie calls his 'standard basis'. The latter is, in many respects, a probabilistic mixture of a number of my specific scenarios and will, accordingly, exhibit higher variability than my 'central' scenario.

12.3 *Extreme Values*

12.3.1 In using the phrase 'far too tame', I also had in mind the fact that extreme values of many financial and economic series tend to be bunched into small periods of time rather than, as with most linear models, being randomly spread over time.

12.3.2 The most important practical consequence is that real world distributions have fatter 'downside' tails than most linear time series fitted by conventional statistical techniques. As a consequence, the underlying level of risk may be understated to a very dangerous extent.

12.4 *Confidence*

12.4.1 It is very difficult for the user of a 'conventional' stochastic model to have confidence in applying the results to real-world situations if, as is usually the case, the model is formulated in highly technical and statistical terms, rather than being transparent in the sense that all the underlying linkages are described in something much closer to plain English to facilitate verification of the appropriateness of the model. My Robinson Crusoe example shows how this problem can be mitigated to some extent by translating the initial statistical formulation into something more recognisable to numerate professionals.

12.4.2 It was clearly this difficulty that led the then President of the Institute to comment, as follows, at the end of the discussion on Geoghegan *et al.* (1992):

> "While I agree with the conclusions of the paper, I believe that, if reliance is to be placed upon the model, this must necessarily be an area of specialisation for a few actuaries rather than trying to take actuaries, generally, deeply into this field."

12.5 *Non-Stationarity*

12.5.1 The apparently intractable problems caused by possible non-stationarity are described very clearly in Geoghegan *et al.* (1992):

> "It is, perhaps, worth introducing this section by emphasising that stationarity is often the most important limitation of economic and financial time-series models, in that they are unable to anticipate future changes in the process which generated historic data."

> "Considering the model under review, there is an implicit assumption that the parameters of the model remain constant over time. However, there is considerable evidence to suggest that the process generating the data is changing over time."

> "However, it is difficult to see how one can make any forecast of unpredictable changes, nor is it obvious how the process might change over time in any predictable way."

12.5.2 This situation is reminiscent, in many ways, of Tolstoy's analogy of Achilles and the tortoise. Just as the mathematical tools available to the Ancient Greeks were unable to solve this paradox, so the mathematical toolkit of the present day financial economist is unable to cope with the very important real world phenomenon of non-stationarity.

12.5.3 Again, the then President commented on the immense practical significance of this issue:

> "In projecting the past, we need to be sure that nothing fundamental has happened or is happening which represents a significant one-off change. The economic events of the last 2 years lead me to wonder if we are not going through a major discontinuity at the present time."

12.6 *General Properties*

12.6.1 Unlike the situation for gilts and equities, where, in each case, two partial differential equations define the general structure of the model, inflation is such a complex and erratic phenomenon that no universally applicable parametric equations are likely to exist. I therefore begin with a 'general reasoning' approach, and describe the essential features of an inflation series that have to be translated into mathematical language.

12.6.2 Inflation is, essentially, a symptom of instability in an economic system. Because of various linkages with wages and prices, negative inflation rates are uncommon, but instability as a result of background economic conditions (e.g. lax monetary policy) or external effects (e.g. commodity price shocks such as the first and second 'oil shocks') can result in a marked rise in the inflation rate above the intrinsic rate. When the rate of inflation is low and has been relatively stable for a number of years, the year-on-year variability is significantly lower than when recent rates have been high.

12.6.3 We can now postulate three underlying mechanisms which affect the year-on-year inflation rate i_t:

(1) a tendency for the rate to return to some 'intrinsic' value;
(2) a random error component which operates every year; and
(3) an upward-only random shock component which operates infrequently.

12.7 *General Formulation*

12.7.1 These general properties can be encapsulated in one line:

Annual change = Centralising effect + Random error + Random shock.

This suggests the following general model:

$$i_t = i_{t-1} - A_t\,(i_{t-1} - B) + D_t + E_t$$

where:

$A_t > 0$;
B is the 'intrinsic' rate;
D_t is the random error term; and
E_t is the random shock term.

12.7.2 A constant value for A_t would imply a constant rate of return to the intrinsic rate, which is highly implausible given that background economic conditions, which may have a significant impact on the rate of inflation, change relatively slowly over time. To provide a mechanism for the inflation rate rising very significantly above the intrinsic rate as a result of general economic conditions, it is essential to use a lower value of A_t when there is an upward trend in inflation. This can be achieved by replacing the term $- A_t\,(i_{t-1} - B)$ by $- A\,(i_{t-1} - B) + C.\,\text{Trend}^{+}$, where A and C are constants and Trend^{+} is the recent trend rate of change of inflation when this is positive and zero otherwise.

12.7.3 To allow for the higher variability of inflation when recent rates have been high, the random error term can be expressed as $D(t-1).Z(t)$ where

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

$D(t-1)$ is positive and increases with the recent average rate of inflation, and $Z(t)$ is a random unit normal variate. Also, since the centralising and random error terms provide the mechanism for a large positive impulse when inflation is high, but cannot provide the mechanism for a large positive impulse when inflation is low, the random shock term can be regarded as upward only.

12.7.4 The general formulation of the model is now:

Annual change = Linear centralising effect + Non-linear trend + Non-linear shock + Non-linear error.

In terms of my different kinds of uncertainty, the trend and shock terms relate primarily to the 'first kind' (i.e. the general level), the error term is clearly recognisable as being of the 'second kind', and the trend term relates primarily to the 'third kind' (i.e. the tendency for high values to be closely bunched together in time).

12.8 *Choice of Auxiliary Functions*

To keep both the formulation of the model and the necessary computations as simple as possible, I chose very simple functions for each of the three non-linear terms. Thus, the trend term uses a geometric moving average, the shock term is of a fixed amount with a specified probability of occurrence, and the error term is essentially a unit normal variate multiplied by a geometric moving average of recent rates of inflation.

12.9 *Choice of Parameter Values*

12.9.1 Since the Wilkie inflation model can be written schematically as:

Annual change = Linear centralising effect + Linear random error

it can be regarded as a very special case of my model with all three non-linear components suppressed; the trend and shock terms are omitted, and the error term has constant variance rather than varying with the recent rate of inflation.

12.9.2 If we regard the Wilkie model as the best available linear model, it is obvious that the value of A in the centralising term of my model should be somewhat larger than the value of 0.4 for the Wilkie model, and, after a little experimentation, I chose 0.5. For the other parameters, I used 'actuarial judgement' rather than any conventional statistical approach.

12.10 *U.K. Inflation from 1951 to 1982*

12.10.1 Since my main criticism of the Wilkie inflation model is that it is 'too tame', in the sense that it does not generate high values that are closely bunched over time, the 'acid test' of my non-linear model is whether, when fitted to past data, the normal variate percentile points are more random over time than is the case for the Wilkie model.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

12.10.2 The two sets of percentile points for 1951 to 1982 are set out in Appendices 1 and 2 of Clarkson (1991). In the case of the Wilkie model, 3 of the 4 top decile values occur in the 4-year period 1974 to 1977, which strongly suggests that the percentile values are not random over time. Also, there is a run of 22 values, from 1952 to 1973, where the residual is less than one standard deviation in absolute magnitude, which is exceptionally strong evidence that the residual is not a random normal variate. In the case of my non-linear model, the residuals are much closer to being randomly distributed over time.

12.10.3 My conclusion that the non-linear formulation gives a better fit is mirrored by the comments of the Wilkie Review Group in Geoghegan *et al.* (1992):

> "This model, which relies very much on actuarial judgement rather than being formulated within a conventional statistical framework, provides an improved fit for post-war U.K. inflation data and results in significantly fewer negative values than is the case with the Wilkie inflation model."

12.11 *Future Projections*

12.11.1 The parameter values that I chose were designed to reflect the 'average' inflation background over the past 30 years or so, but, because of the complex interactions between the various non-linear components of my model, I could not specify directly a mean rate of inflation. I therefore ran the model for 200 years (using pen, paper and a table of random numbers rather than a computer) to see what mean annual rate of inflation resulted. The outcome of 6.00% seemed eminently satisfactory, and is an interesting antithesis to the conventional methodology of the financial economist, who uses the mean and variance of the data to fit some standard model rather than attempting to identify, and then calibrate, each of the underlying causal mechanisms.

12.11.2 I now return to the problematic area of non-stationarity. If we assume that the general formulation of my non-linear model is broadly satisfactory, and that the specific parameter values used give an adequate representation of the past, then projecting the future reduces to assessing what specific parameter values will be appropriate given:

(1) our understanding of past economic conditions;
(2) our understanding of present economic conditions and, in particular, of what structural changes have occurred; and
(3) our judgement as to what future structural changes are likely.

12.11.3 Expressed in these terms, stochastic modelling of future inflation scenarios is no different, in principle, from the assessment of future mortality rates, which involves three quite distinct stages:

(1) the production of a standard mortality table;
(2) a review (either CMI Committee or internal) of the most recent experience relative to the standard table; and
(3) actuarial judgement as to what modifications to this recent experience are appropriate in the specific circumstances of the situation.

12.12 *Further Comparisons with the Wilkie Model*

12.12.1 In Clarkson (1991) I describe two further diagnostic tests as to whether the additional complexity of my non-linear model is justified by a significantly more realistic portrayal of the post-war U.K. inflation data. In both cases my answer is in the affirmative.

12.12.2 Firstly, I show by 'general reasoning' that, in the absence of my non-linear components, the percentiles of the normal variate corresponding to the residuals will exhibit a frequency distribution that is far from random. First and third 'quarter' values will be significantly below and above the average, respectively, while second and fourth 'quarter' values will be above and below the average, respectively. The results using the percentile values in Appendices 1 and 2 show precisely this behaviour for the Wilkie linear model, but a pattern very close to random for my non-linear model:

	Frequency	
Percentile value	Wilkie (1986)	Clarkson (1991)
1– 25	3	7
26– 50	10	8
51– 75	12	8
76–100	7	9

12.12.3 My second test is elementary in the extreme. I express the absolute value of the residual in the Wilkie model as a multiple of my geometric average value of the force of inflation, and then classify this multiple as either 'low', 'medium', 'high', or 'very high'. Taking 8 bands for the average force of inflation, I obtain a 4 x 8 matrix whose elements will be randomly distributed if the expected value of the residuals is independent of the recent rate of inflation. If, on the other hand, the expected value of the residual is — as I postulate — directly proportional to the recent rate of inflation, the matrix elements will cluster along a straight line from top left to bottom right. The resulting matrix is exceptionally strong evidence in support of my postulate:

1	2	5	2	1	0	0	1
0	2	2	1	0	1	1	2
0	1	0	0	1	0	2	2
0	0	0	0	1	0	0	4

12.12.4 Although my non-linear model may, at first sight, appear somewhat complex, the computations involved are quite straightforward. The model can be run 'longhand' line by line using a schedule with 14 columns, and, indeed, the full working for my simulation over 200 years required no more than 5 pages of A4, i.e. 40 lines of calculation per page. In computer terms, the computations are trivial, and accordingly, when considering whether or not the use of my non-linear model is appropriate, the computational practicalities should not be an issue.

12.13 *Engle and Heteroscedasticity*

12.13.1 Because of length constraints (AFIR papers are generally restricted to a maximum of around 20 pages), I was able to include in Clarkson (1991) only one paragraph on Engle's pioneering 1982 paper in *Econometrica*:

> "An increasing awareness by statisticians in recent years of the limitations of linear models as representations of many economic and financial series has led Engle (1982) and others to develop autoregressive conditional heteroscedastic (ARCH) models to remove the implausible assumption of a constant variance. However, as with ARIMA models, the statistical tests only address the question of whether a particular model is the best of its class rather than addressing the more fundamental question of whether any model of that class offers a satisfactory representation. Also, asymmetric non-linear components such as the upward-only trend component in my non-linear model cannot be incorporated."

In view of the title and objectives of the present paper, it is clearly essential to comment in much more detail on Engle's work on heteroscedasticity in general and on its application to projections of U.K. inflation in particular.

12.13.2 The first paragraph of Engle's summary of his paper is as follows:

> "Traditional econometric models assume a constant one-period forecast variance. To generalise this implausible assumption, a new class of stochastic process called autoregressive conditional heteroscedastic (ARCH) processes are introduced in this paper. These are mean zero, serially uncorrelated processes with nonconstant variances conditional on the past, but constant unconditional variances. For such processes, the recent past gives information about the one-period forecast variance."

In plain English, and in the context of a price inflation index, such as the RPI in the U.K., this says that the higher the annual rate of inflation the higher will be the year-on-year variability. A glance at the history of changes in the RPI over calendar years from, say, 1950 to date, shows that this is obvious.

12.13.3 After developing the theory behind his ARCH regression model, Engle suggests that it has a variety of characteristics which make it attractive for econometric applications. Firstly, he observes that econometric forecasters have found that their ability to predict the future varies from one period to another, and he cites two observations which mirror some of my earlier comments:

> "The inherent uncertainty or randomness associated with different forecast periods seems to vary widely over time"

and

> "Large and small errors tend to cluster together (in contiguous time periods)."

12.13.4 The second example, from monetary theory and the theory of finance, shows the crucial importance that financial economists attach to expected means and variances:

> "By the simplest assumptions, portfolios of financial assets are held as functions of the expected means and variances of the rates of return. Any shifts in asset demand must be associated with changes in expected means and variances of the rates of return."

12.13.5 The third example is, perhaps, the most interesting of all, in that Engle suggests that detection of an ARCH effect should often be interpreted as evidence of mis-specification through making 'standard assumptions about the disturbances':

> "A third interpretation is that the ARCH regression model is an approximation to a more complex regression which has non-ARCH disturbances. The ARCH specification might then be picking up the effect of variables omitted from the estimated model. The existence of an ARCH effect would be interpreted as evidence of misspecification, either by omitted variables or through structural change. If this is the case, ARCH may be a better approximation to reality than making standard assumptions about the disturbances, but trying to find the omitted variable or determine the nature of the structural change would be even better."

12.13.6 It might have been thought that Engle's non-linear model for U.K. inflation, which is constructed from a financial economist's viewpoint, could be compared with the linear actuarial model in Wilkie (1986) and the non-linear actuarial model in Clarkson (1991). Such, however, is not the case. Engle's 'model' for U.K. price inflation as the dependent variable uses U.K. wage inflation as the independent variable:

> "A conventional price equation was estimated using British data from 1958-II through 1977-II. It was assumed that price inflation followed wage increases; thus the model is a restricted transfer function."

Engle only studies relativities between two similar inflation series, rather than projecting future values, as is required in actuarial applications.

12.13.7 Although he assesses the properties of U.K. price inflation in relative, rather than absolute, terms (in that the properties of U.K. wage inflation are not discussed), Engle's comments on the non-randomness of residuals over time are remarkably similar to the corresponding comments in ¶12.10.2:

> "In order to determine whether the confidence intervals arising from the ARCH model were superior to the least squares model, the outliers were examined. The expected number of residuals exceeding two (conditional) standard deviations is 3.5. For ordinary least squares, there were 5 while ARCH produced 3. For least squares these occurred in '74-I, '75-I, '75-II, '75-IV, and '76-II; they all occur within three years of each other and, in fact, three of them are in the same year. For the ARCH model they are much more spread out and only one of the least squares points remains an outlier, although the others are still large. Examining the observations exceeding one standard deviation shows similar effects."

12.13.8 Engle concludes from this analysis that modelling the ARCH effect leads to a better fit (i.e. 'more truly random' residuals in terms of time of occurrence as well as frequency):

> "Thus the number of outliers for ordinary least squares is reasonable; however, the timing of their occurrence is far from random. The ARCH model comes closer to truly random residuals after standardizing for their conditional distributions."

12.14 *Criteria for Stochastic Models*

12.14.1 Since it is obvious that my bespoke non-linear model cannot be assessed on the basis of any of the conventional statistical tests used by financial economists, I suggest, in Clarkson (1991), that it would be useful to have general criteria for the evaluation of stochastic models, whether linear or non-linear. I also have in mind the fact that the very comprehensive estimation procedures and statistical tests in Wilkie (1984) may well identify the 'best' ARIMA model for U.K. inflation, but nevertheless cannot provide a complete answer to the question 'Is an ARIMA model appropriate?'.

12.14.2 The criteria that I suggest are:

(1) The mechanisms implied by the model should be plausible on economic grounds.
(2) The parameter values and auxiliary functions should be such that, in fitting the model to the past experience, the standardised residuals (e.g. the residuals divided by $D(t-1)$ in the case of my non-linear inflation model) should be consistent with a random unit normal variate, not only in terms of frequency, but also in terms of order.
(3) The parameter values and auxiliary functions should be plausible on economic grounds.

12.14.3 The second criterion is quite different from, and much stronger than, most conventional statistical tests, which generally assume that the residuals are normal in frequency, but ignore the question of order.

12.15 *Stockmarket Efficiency*

12.15.1 I also suggest, in Clarkson (1991), that what I regard as the misguided pronouncements by financial economists as to the 'efficiency' of capital markets were not unrelated to the very primitive nature of the linear models that Jensen and others used to 'test' for efficiency. I shall address the efficiency paradox in more detail later, but in the context of the debate between linearity and non-linearity, two quotations from the academic literature are of particular relevance.

12.15.2 Firstly, Mandelbrot (1963) comments as follows in the context of a capital market model which gives fatter tails than is the case for standard distributions:

> "Broadly speaking, the predictions of my main model seem to me to be reasonable. At closer inspection, however, one notes that large price changes are not isolated between periods of slow change; they rather tend to be the result of several fluctuations, some of which "overshoot" the final change. Similarly, the movement in prices in periods of tranquillity seem to be smoother than predicted by my process. In other words, large changes tend to be followed by large changes — of either sign — and small changes tend to be followed by small changes."

These observations reflect, very closely, various comments of mine and of Engle in the context of the non-linear nature of inflation.

12.15.3 Secondly, in his seminal review work on stockmarket efficiency, Fama (1970) describes how non-normal stable distributions of precisely the type advocated by Mandelbrot are, in many ways, more realistic than standard distributions, but then observes somewhat ruefully:

> "Economists have, however, been reluctant to accept these results, primarily because of the wealth of statistical techniques available for dealing with normal variables and the relative paucity of such techniques for non-normal stable variables."

Tolstoy's analogy of the unsupervised plasterers comes readily to mind; Fama is suggesting that economists are opting for a spuriously scientific 'second best' at the expense of real-world relevance.

12.16 *Conclusions*

The marked non-linearity of the U.K. inflation series restricts the applicability of linear stochastic models such as the Wilkie model. The incorporation of an ARCH or similar effect, while perhaps a small step in the right direction, is unlikely to be a radical enough response to observed departures from linearity.

13. Some Observations on the Theory of Games

13.1 *Introduction*

Since the mathematical framework that has become known as the theory of games is the foundation work within economic science for taking into account 'the actions of others', I investigated the theoretical and practical relevance of the theory of games, and set out my conclusions in Clarkson (1993).

13.2 *A Sherlock Holmes Paradigm*

13.2.1 A crucial theoretical construct in von Neumann & Morgenstern (1944) is the 'two-person zero sum game', which the authors say is the appropriate underlying theory for determining optimal behaviour in many instances of real world economic life. The application of their theory is illustrated by a case study involving part of the narrative of *The Final Problem* by Sir Arthur Conan Doyle; my summary of the background, as cited in Clarkson (1993), is given in the following paragraph.

13.2.2 In an episode from *The Adventures of Sherlock Holmes* by Sir Arthur Conan Doyle, Holmes, hotly pursued by Professor Moriarty, has boarded a train at Victoria Station, London, in an attempt to reach Dover and escape to the Continent. The only intermediate station is Canterbury. In essence, Holmes and Moriarty have independently to decide at which station to detrain. If both detrain at the same station, Holmes will be killed by Moriarty. If Holmes detrains at Dover while Moriarty detrains at Canterbury, Holmes survives and can escape to the Continent. This outcome is more advantageous to Holmes than if he detrains at Canterbury while Moriarty detrains at Dover, in which case Holmes escapes for the moment, but does not reach the Continent.

13.2.3 Using highly arbitrary utility values to Moriarty of 100 if they both detrain at either Canterbury or Dover, −50 if Moriarty detrains at Canterbury and Holmes continues on to Dover, and 0 if Holmes detrains at Canterbury and

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

Moriarty continues on to Dover, von Neumann & Morgenstern suggest that Moriarty should detrain at Canterbury with a 40% probability and detrain at Dover with a 60% probability. Taking utility to Holmes as being the negative value of utility to Moriarty, the optimal course of action suggested for Holmes is to detrain at Canterbury with a 60% probability and to detrain at Dover with a 40% probability. If these mixed strategies are followed, there is a probability of 0.48 that Moriarty and Holmes detrain at the same station. Accordingly, von Neumann & Morgenstern conclude that "Sherlock Holmes is as good as 48% dead when his train pulls out from Victoria Station."

13.2.4 In common sense terms, I see four very serious objections to this solution. First, reflective real world behaviour tends to be characterised by difficult either/or decisions involving a balance of probabilities rather than by compromise probability mixtures; the mixed strategy approach is, therefore, unnatural in the extreme. If the ten directors of an insurance company have to choose between Jones and Smith as the new Chief Executive, with six favouring Jones and four favouring Smith after all reasonable opportunity for debate, they appoint Jones and have this minuted as a collective decision. They do not carry out a lottery by drawing from a hat containing ten slips of paper, six marked Jones and four marked Smith. Second, the crucial 'zero sum' assumption that Holmes and Moriarty have symmetric utilities is untenable from a psychological point of view. Failure to Holmes means immediate death, whereas, to Moriarty, Holmes' death would be only one more in a long list of murders in the furtherance of a criminal career. Third, Holmes can deduce from Moriarty's strategy that his own probability of being killed has a minimum value of 40% if he detrains at Canterbury with certainty. Since the benefit of being at Dover rather than Canterbury should Moriarty leave at the other station pales into total insignificance compared to the 8% increase in his probability of survival if he leaves the train with certainty at Canterbury, Holmes will, of course, choose Canterbury, and this is, indeed, how Conan Doyle's narrative proceeds. Fourth, the assumption that Holmes will be killed with certainty should both he and Moriarty detrain at the same station is fatalistic in the extreme. Holmes and Watson could, for instance, leap from the train, run out of the station and jump into a hansom cab. As a refinement, Holmes could afterwards leave Watson alone in the cab as a decoy, and either pick up another cab or take refuge temporarily in a shop or other building. The branching possibilities are virtually limitless.

13.3 *More on the Sherlock Holmes Paradigm*

13.3.1 I had assumed that the account of this episode in von Neumann & Morgenstern (1944) would be an accurate reflection of the situation as narrated by Conan Doyle, and, accordingly, in writing Clarkson (1993), I used the von Neumann & Morgenstern account without referring back to the original narrative. On reading the relevant narrative for another purpose a year after the presentation of my paper, I was surprised, and more than a little aggrieved, to find that the von Neumann & Morgenstern version deviates in many material respects from Conan Doyle's description of events.

13.3.2 Von Neumann & Morgenstern categorically state that Moriarty, on the platform at Victoria Station, has seen Holmes on the train pulling out. According to the narrative, Holmes could see Moriarty amidst a crowd of people some distance down the platform. However, even had Holmes not been still disguised as a priest, it would have been impossible for Moriarty to have recognised him inside the compartment from such an oblique angle. Moriarty may have inferred from Watson's arrival at the station that it was highly probable that Holmes was on the train, but that is quite different from certain knowledge. Also, Holmes' disguise, which fooled even his trusted friend Watson at close quarters, would have made it even easier to elude Moriarty should both have chosen to detrain at the same station. But perhaps the most blatant distortion relates to the success, or otherwise, of Holmes in achieving his chosen objective. It is clear from the original narrative that Holmes has a very definite objective in mind, namely to stay alive until the following Monday, when a massive police operation can be launched to round up Moriarty and all the others in his criminal network. By taking the next train from Canterbury to Newhaven, Holmes and Watson are not only able to reach the Continent that day, but they have also succeeded in shaking Moriarty and his gang off their trail. Far from being a less advantageous outcome, as stated by von Neumann & Morgenstern, escaping to the Continent via Canterbury and Newhaven is of far greater advantage to Holmes than reaching Dover before Moriarty and thereafter sailing to France.

13.4 *Complexity, Freewill and the Dimension of Time*

13.4.1 The above considerations lead me to the general conclusion that the theory of games approach is fatally flawed as a framework for 'intelligent' or 'rational' behaviour, in that it suppresses, rather than enhances, our innate capabilities in three crucial areas of human existence — complexity, freewill, and the dimension of time.

13.4.2 While we may suppress the complexity of our knowledge base by pretending, for instance, that Moriarty and Holmes have seen each other, real life involves 'shades of grey', rather than each and every datum of information being either 'known' or 'unknown' to each and every participant.

13.4.3 While we may suppress the freewill of the individual by pretending, for instance, that Holmes will stand like a tailor's dummy on whatever station platform his unthinking drawing of a lottery ticket tells him to alight at, influential decision-makers in the real world are well aware of the 'lateral thinking' and 'brainstorming' capabilities of the human mind to identify plausible courses of action that might be of assistance in the pursuit of some specified goal.

13.4.4 While we may suppress the dimension of time by pretending that the end of some arbitrary time period, such as the arrival of the boat-train at Dover, is the point in time at which to assess 'success' or 'failure', any meaningful

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

measure of success in the real world must renounce this 'short-termism', and incorporate, instead, some appropriate longer-term time scale.

13.5 *Conclusions*

By suppressing crucial elements of complexity, freewill and the dimension of time, the theory of games approach would appear to hinder, rather than help, our quest for a better framework for human behaviour than that provided by current theories of financial economics. The theoretical foundations of the theory of games are discussed in more detail in Section 18.

14. THE COMING REVOLUTION IN THE THEORY OF FINANCE

Clarkson (1994) is essentially an outline sketch of some of the material presented in this paper, in that it discusses unsatisfactory features of some key elements of modern finance theory such as the Capital Asset Pricing Model and utility theory. In particular, I show that the Allais Paradox, which I regard as the utility theory equivalent of the paradox of Achilles and the tortoise, can be resolved very easily within my new downside framework for risk.

15. SOME OBSERVATIONS ON THE BLACK-SCHOLES OPTION PRICING FORMULA

15.1 *Simplifying Assumptions*

15.1.1 The mathematics of option pricing and of dynamic hedging strategies using options and other derivatives is largely based on the Black-Scholes option pricing formula, which assumes that a risk-free hedge can be set up, at any instant in time, without affecting the underlying market prices. In Clarkson (1995a) I suggest that this assumption is not only highly unrealistic, but also likely to create a false sense of security amongst investors or financial institutions who think that they are reducing risk through dynamic hedging techniques.

15.1.2 A very clear account of the resulting potential for financial accidents is given by George Soros in his evidence to the U.S. House Banking Committee in April 1994:

> "The trouble with derivative instruments is that those who issue them usually protect themselves against losses by engaging in so-called delta, or dynamic hedging. Dynamic hedging means, in effect, that if the market moves against the issuer, the issuer is forced to move in the same direction ... and thereby amplify the initial price disturbance. As long as price changes are continuous, no great harm is done except perhaps higher volatility, which, in turn, increases demand for derivative instruments. But if there is an overwhelming amount of dynamic hedging to be done in the same direction, price movements may become discontinuous. This raises the specter of financial dislocation. Those who need to engage in dynamic hedging, but cannot execute their orders, may suffer catastrophic losses .... That is what happened in the stock market crash of 1987."

15.2 *Conclusions*

Trading strategies, based on the continuous rebalancing assumption, will often increase, not decrease, the level of financial risk. A better theory of option pricing, built up from less unrealistic simplifying assumptions, could greatly reduce the general level of systemic financial risk.

16. ASSET-LIABILITY MODELLING AND THE DOWNSIDE APPROACH TO RISK

16.1 *Introduction*

Clarkson (1995b) was written to investigate the application of my downside framework for risk to asset-liability modelling.

16.2 *Market Yields*

The main insight, as far as asset values is concerned, is that equity market levels, suitably measured, can be expected to oscillate around fixed-interest market levels in a manner not unlike that of my Dynamic Equilibrium Model. While based on practical experience, this conjecture is strongly supported by the results of co-integration vector analysis, as set out in Mills (1991). The most obvious place to look for stable patterns of behaviour is in what used to be known as the 'confidence indicator', namely the yield ratio of a fixed-interest yield to the dividend yield on the equity market. An important part of the description of future asset values is then given by the probability distribution of the fixed-interest yield and, for each value of this fixed-interest yield, the probability distribution of the yield ratio.

16.3 *Scenario Analysis*

For a certain set of specified values of the rate of real dividend growth and the inflation rate, consistent values for the assets and liabilities can be calculated directly. By expressing this result as a solvency ratio (e.g. 92%) of the value of the assets to the value of the liabilities, we obtain an immediate and meaningful measure of downside risk. Different sets of specified values of the economic variables, corresponding to most likely and then higher and lower values, can then be analysed to build up the general picture and to identify which combinations of economic conditions and market yields pose the greatest threat to ongoing solvency. The analysis can then be repeated for different distributions of assets, to determine whether a change of investment policy is appropriate, either to reduce the risk of insolvency or, where the solvency position is satisfactory, to increase the likely return on assets.

16.4 *Conclusions*

Many actuaries have used mean-variance optimisation techniques in asset-liability modelling applications. The alternative approach, outlined above, is far more direct, far easier to understand and to explain to clients, and — I believe — potentially far more powerful.

17. Utility Theory

17.1 *Introduction*

In Section 11, using the framework of Clarkson (1990), I conclude that utility theory, both as originally propounded by Bernoulli in 1738 and also as presented in a modern context in the actuarial textbook Bowers *et al.* (1986), can lead to conclusions that are inconsistent with common sense approaches, and must, therefore, be rejected as a framework for rational choice under conditions of risk and uncertainty on grounds of practical relevance. In this section I examine the mathematical foundations of utility theory.

17.2 *The von Neumann & Morgenstern Axioms and the Dimension of Time*

17.2.1 Modern interest in utility theory essentially began when von Neumann & Morgenstern (1944) proved mathematically that, if a system of utilities existed and satisfied a set of apparently innocuous axioms, then the utility measure is unique up to a linear transformation. This result is what I call the fundamental postulate of utility theory, namely that a utility function, unique up to a linear transformation, can be used to calculate expected utility as a ranking measure of the attractiveness of different probabilistic outcomes.

17.2.2 The last of their axioms, namely (3:C:b), is the statement that it is irrelevant whether a combination of two constituents is obtained in two successive steps using first the probabilities a and $(1-a)$ and then the probabilities b and $(1-b)$, or in one operation using the probabilities c and $(1-c)$ where $c = ab$.

17.2.3 Since this is the utility theory equivalent of the multiplication axiom for independent events within probability theory, it appears, from a mathematical and philosophical viewpoint, to be uncontroversial in the extreme. However, in Section 10 I show how this independence axiom breaks down completely in the case of probability mixtures of sports, which, in themselves, involve probability mixtures of unwanted outcomes. There is a point of discontinuity in the human mind's capacity to deal with increasing levels of risk, and above this level deliberative thought is impossible. In mathematical terms, the assumption of independence breaks down, in that the rate at which the human mind can absorb and dissipate increasing levels of physical risk collapses to zero at a certain shear point rather than remaining constant. Given the close correspondence between physical risk and financial risk, as described in Section 10, I conclude that a similar psychological shear point will exist for financial risk, that Axiom (3:C:b) is untenable, and that the existence proof for a utility function (as generally understood) is invalid.

17.2.4 A physical analogy, which similarly introduces the dimension of time, is someone descending from the second floor level in a building to ground level. Walking down two flights of stairs will take around fifteen seconds and involve negligible risk. Jumping the sixteen feet or so to the ground will take only one second, but is likely to cause serious physical injury. Over fifteen seconds the human frame can absorb, without difficulty, the same total amount of kinetic energy that, if compressed into one second, would cause very serious damage.

17.2.5 It is interesting to note that von Neumann & Morgenstern themselves express considerable reservations about the validity of Axiom (3:C:b). In the main text they suggest that it should be retained "unless a much more refined system of psychology is used than the one now available for the purposes of economics". At the end of the Appendix added to the second edition they comment as follows:

> "It seems probable that the really critical group of axioms is (3:C), or, more specifically, the axiom (3:C:b) .... Some change of the system (3:A) — (3:C), at any rate involving the abandonment or at least a radical modification of (3:C:b), may perhaps lead to a mathematically complete and satisfactory calculus of utilities .... It is hoped that a way will be found to achieve this, but the mathematical difficulties seem to be considerable."

17.2.6 I am in no doubt that the theoretical foundations of rational choice under uncertainty and risk have to be rebuilt without the independence axiom (3:C:b); my suggested framework in Section 10 presents no mathematical difficulties.

17.3 *The Savage Axioms*

17.3.1 While the axiomatic approach to utility theory, pioneered by von Neumann & Morgenstern, was seen as an exceptionally important breakthrough in economic science, many prominent economists, such as Friedman, expressed serious reservations about the validity of the axioms, with the so-called 'independence axiom' attracting particularly severe criticism. Various mathematicians proposed alternative axiomatic frameworks which, while similar in terms of their mathematical implications, could be described in plain English. The approach which found most favour with statisticians and mathematical economists is that described in Savage (1954), a modern commentary on which can be found in Anand (1993).

17.3.2 My first observation is that, since no fewer than eight axioms, some of a highly technical nature, are required, it is completely beyond the grasp of the human mind to translate the logic of these axioms into practical action. My 'Who is Prime Minister?' logic puzzle is trivial by comparison.

17.3.3 It may be suggested by proponents of utility theory that the internal consistency of the Savage axioms provides a valuable framework for deliberative analysis of complex real world problems, and that, given a sufficiently powerful computer, meaningful practical solutions can be generated. I disagree; two axioms, in particular, are, in my view, fatally flawed as building blocks for any intelligent approach to human choice under uncertainty and risk.

17.3.4 Axiom 5 relating to 'Expected Wealth Independence' is:

> "If a decision-maker is about to place a stake on one of two lotteries, the decision should be based on the probability of winning but not on the amount of the stake."

If we increase the size of the stake to a value which would prejudice personal or corporate survival in the event of an adverse outcome, it is apparent that this axiom is an exhortation to pursue perceived short-term advantage without any reasoned consideration of the consequences of failure. I shall use this as a descriptor of one crucial element of what I call 'unintelligent' or type-B behaviour.

17.3.5 Axiom 7 relating to 'Continuity in Probability' and Anand's commentary thereon are:

> "There exists a state with a strictly positive probability so small that any change to the pay-off in that state, however large, will not affect the preference for that act. This is, in effect, a decision-theoretic parallel to the statistical convention that very unlikely events should be regarded as having zero probability."

Avalanches in ski-ing, major natural disasters, and exceptional capital market movements are all examples of very unlikely events. To suggest that they should all be given zero probability is so utterly ridiculous from a common sense viewpoint that I shall use this axiom as the cornerstone for my formulation of 'unintelligent' or type-B behaviour.

17.4 *Tversky and Risk Aversion*

17.4.1 The generally accepted academic formulations of 'increasing risk', as described in Rothschild & Stiglitz (1970, 1971), are totally impenetrable except to expert mathematicians. However, as might be expected, numerous economists and psychologists have shown that observed real world behaviour is often inconsistent with these theoretical teachings. In particular, Tversky (1978) questions the fundamental assumption that individuals are 'risk averse' in the generally accepted sense. In the following description of his best known example I have taken the liberty of translating U.S. dollars into British pounds.

17.4.2 An investor has the choice between a certain gain of £85,000 or an 85% chance of receiving £100,000 and a 15% chance of receiving nothing. Most people will prefer the certain outcome, which is consistent with theory which suggests that individuals are risk-averse. Suppose now that the investor has a choice between a certain loss of £85,000 or an 85% chance of losing £100,000 and a 15% chance of losing nothing. In this situation people will 'gamble' and choose the latter, despite the significant chance of further loss. Since such behaviour is inconsistent with the assumption of individuals being 'risk-averse', the theory breaks down.

17.4.3 Such behaviour is consistent with my alternative framework for uncertainty and risk, which incorporates a specific personal threshold of risk.

17.4.4 Having detected anomalous behaviour in the context of the existing

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

paradigms of financial economics, Tversky then questions the 'rational expectations' assumption that rational investors can formulate beliefs and subjective probabilities that are precise and unbiased. In particular, he discusses the general bias towards overconfidence (precisely the same 'absurd presumption of their own good fortune' described by Adam Smith), and suggests how such behaviour, while perhaps a desirable feature in survival terms, may cause individuals to fail to take into account information that is available to others.

17.5 *A Conference in Toulouse*

17.5.1 In September 1994 I attended a conference on utility theory in Toulouse, organised under the auspices of the Geneva Association, an objective of which is to support and encourage research in the economics of risk, uncertainty, insurance and related issues. Since the keynote speaker was the U.S. economist Machina, I studied in some detail the areas of so-called anomalous behaviour described in Machina (1982), and came to the conclusion that nearly all the examples of real world behaviour that were regarded as inconsistent with the expected utility paradigm were precisely the types of behaviour that would be expected under my alternative framework, as described in Section 10.

17.5.2 During the conference I detected two main themes behind the presentations, namely a willingness, in the light of perceived 'probability biases', to abandon probability theory as the cornerstone of rational choice under uncertainty and the acceptance of a fragmented, rather than unified, approach, whereby more and more complex axiomatic models are put forward to 'explain' specific instances of perceived anomalous behaviour. Both intuitively and by analogy with the history of astronomy, I regard both of these developments as steps backwards, rather than forwards, in our understanding of human choice under uncertainty. I describe, below, my alternative explanations of the phenomena known to mathematical economists as 'probability biases' and 'framing and coding effects'.

17.6 *Donald's Paradox and Framing*

17.6.1 The 'probability bias' that I discuss is a variation of the 'How many people in a room?' paradox, often cited by the late D.W.A. Donald.

17.6.2 There are 35 people in a room. Assuming a uniform distribution of birthdays over the calendar year, is the probability that at least two people have the same birthday:
(1) around 0.1;
(2) around 0.4; or
(3) around 0.8?

17.6.3 To an actuary, it is quite straightforward to determine that the probability is around 0.8. The element of paradox is that most people, even other numerate professionals, will opt for 'around 0.1', and express incredulity when told the correct answer.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

17.6.4 My first observation is the obvious one that mental calculations of other than a trivial nature are generally outwith the capabilities of the untrained human mind. This is consistent with Penrose's refutation of the algorithmic nature of human thought.

17.6.5 My second observation is that the characteristic answer of 'around 0.1' is an example of what I call 'instinctive pattern recognition thought'. We learn from everyday experience that simple proportion is often a very useful approach. Since the required probability varies from zero when there is only one person in the room, to unity when there are 366, applying a simple proportion suggests that when there are 35 the probability will be around 0.1. The situation, however, is grossly non-linear, and the linear approximation is wildly inaccurate.

17.6.6 Another probability-related phenomenon is what economists call 'framing and coding effects.' It is found that an individual, confronted with a choice involving probabilistic situations, will often make different choices depending on how the background information is described or 'framed'. Also, as would be expected on commonsense grounds, an individual's success in making consistent decisions increases as the background is explained in more and more detail as a series of simple choices between alternative scenarios. Economists, however, have interpreted this as evidence that rational choice is based partly on how the information is presented, and have suggested new, and more complex, theoretical approaches, such as prospect theory. Trying to classify observed behaviour as either 'intelligent' or 'unintelligent' would appear a better way forward than this fragmented approach of constructing more and more complex mathematical models for specific aspects of observed behaviour.

17.7 *Parallels with Astronomy*

Since I regard Bernoulli's pioneering work as being one level above the previous pure expectations approach, which itself might be regarded as the rational choice equivalent of the simple, but incomplete, approach of concentric circles in astronomy, I shall place Bernoulli at Level 2. Economists such as Allais, Machina, and Tversky have produced enough evidence about the incompleteness of the Bernoulli approach to be regarded as the equivalents of Copernicus, and thus placed at Level 3. Allais has also, I believe, given sufficiently precise descriptions of the general characteristics of a better approach to justify inclusion at Level 4, corresponding to the achievements of Tycho Brahe, in assembling empirical data, and of Kepler, in identifying specific laws of behaviour. Von Neumann & Morgenstern and Savage, on the other hand, appear to have sacrificed so much of the generality of Bernoulli's approach in the interests of a narrow axiomatic formulation, that I place them at Level 1, the level below Bernoulli.

17.8 *Conclusions*

The main conclusion which I draw is that, despite the excellent guiding principles laid down by Bernoulli more than a quarter of a millenium ago, the theory of rational choice under uncertainty has fallen into a mathematical 'black hole' over the past fifty years, as a result of the dominance of the axiomatic approach to utility theory.

18. THE THEORY OF GAMES

18.1 *Introduction*

While I agree, unreservedly, that it is essential to take into account the behaviour of 'the others' when constructing a formal theoretical framework for any economic or financial system, I do not believe that the mathematical methodologies of the theory of games are appropriate in this regard. I summarise below my reasons for this viewpoint.

18.2 *The Suppression of the Dimension of Time*

18.2.1 To avoid completely the serious difficulties involved in comparing financial events at different points in future time, von Neumann & Morgenstern (1944) assume that all events are located "at one and the same, standardized, moment, preferably in the immediate future" and comment, in a footnote, that it is well known that incorporating the dimension of time "presents very interesting, but as yet extremely obscure, connections with the theory of saving and interest, etc."

18.2.2 The obvious comment here is that the actuarial 'compound interest' approach, as described in Donald (1970) and earlier text-books, has been available for centuries as a highly successful practical methodology.

18.2.3 A less obvious suppression of the dimension of time relates to the assumption that an individual's strategy does not change over time, even if his personal circumstances change materially as the result of earlier rounds of the game. This is inconsistent with Bernoulli's original maxim that behaviour is dependent on the circumstances of the individual.

18.3 *The Suppression of Risk*

Since the 'result' of a game is regarded as its expected outcome with no consideration of the consequences of failure, von Neumann & Morgenstern are effectively turning the clock back to the pure expectations paradigm that Bernoulli sought to replace.

18.4 *The Suppression of Complexity*

The number of distinct outcomes that can be accommodated is very seriously constrained by the necessity to obtain unique solutions to linear equations. In the Sherlock Holmes example, for instance, the suggested solution can be obtained far more easily by equating partial derivatives of expected utility to zero, which is amusingly reminiscent of Tolstoy's comments on the paradox of Achilles and the tortoise. However, if Holmes had even one further option available to him, such as running from Canterbury station rather than waiting fatalistically on the platform, the simplistic methodology of the two person zero-sum game breaks down completely.

18.5 *The Suppression of Freewill*

An immediate corollary to the suppression of complexity is that the freewill of the individual to improve his or her chances of success by, first of all, identifying further plausible options by means of 'lateral thinking' or 'brainstorming', and then analysing their advantages and disadvantages in the light of personal circumstances, is also suppressed. The theory simply cannot cope with these additional possibilities.

18.6 *The Suppression of Uncertainty*

Probabilities for all relevant events are assumed to be known exactly as the limiting values of relative frequencies in long runs of trials. Real life, however, is quite different, as Gray (1967) observes:

> "There are many systems or situations in which repeated trials under identical conditions cannot be made. For example the probability that £100 of $3\frac{1}{2}$% War Loan will increase in value over the next six months cannot be assessed as the limit of relative frequency. Any measure of such probability is only some intuitive degree of belief which may differ considerably from person to person. There is no mathematical method of calculating these probabilities."

18.7 *The Existence of a Competitive Equilibrium*

18.7.1 While it might be thought that the real world relevance of the theory of games is of no practical significance for everyday economic and financial behaviour, the converse is, I believe, the case. In precisely the same way as Bernoulli sought, in his seminal 1738 paper, to replace the pure expectations paradigm with a more realistic framework that took account of the circumstances of the individual, Keynes, in his *General Theory*, sought to replace the linear equilibrium-based paradigm of classical economic theory with a more realistic framework that took account of the observed oscillatory nature of economic events.

18.7.2 Perhaps the most important reason for the decline of the influence of Keynsian thought was the 'proof', in Arrow & Debreu (1954), that a competitive economic system always reaches an equilibrium and that a competitive equilibrium is optimal from a social point of view. Their general proof, however, assumes, crucially, that a competitive economy can be regarded as an abstract game, and, hence, that the mathematical methodologies of the theory of games can be applied. My viewpoint is that this general proof, while of considerable theoretical interest in establishing a benchmark position against which to measure real world behaviour, is invalid as a descriptor of likely real world behaviour, in that the vital aspects of time, risk, uncertainty, complexity and freewill are suppressed.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

18.7.3 Since Arrow and Debreu were awarded Nobel Prizes for Economics in 1972 and 1983 respectively, largely for their contributions to the theory of economic equilibrium, it is not surprising that the paradigm of economic equilibrium has remained dominant in recent times over competing approaches, as propounded, in particular, by the disciples of Keynes.

18.8 *Conclusions*

Having carried out the 'due diligence' and 'verification' investigations that seemed appropriate in the light of my training, both as a mathematician and as an actuary, my conclusion is that we must look elsewhere for a formal framework capable of accommodating the behaviour of 'the others' in economic and financial systems. As I describe in more detail below, I believe that the 'theory of sports' should replace the 'theory of games' as the intermediate formal framework.

19. Portfolio Theory

19.1 *Introduction*

In this section I discuss the development of portfolio theory from its highly theoretical origins in Markowitz (1952) to its highly pragmatic practical implementation, as described in November 1993 by Professor Harry Markowitz, then a Nobel Laureate, at the Royal Society public discussion meeting at which he was the keynote speaker.

19.2 *Markowitz (1952)*

19.2.1 The opening sentences are as follows:

> "The process of selecting a portfolio may be divided into two stages. The first stage starts with observation and experience and ends with beliefs about the future performances of available securities. The second stage starts with the relevant beliefs about future performances and ends with the choice of portfolio. This paper is concerned with the second stage. We first consider the rule that the investor does (or should) maximize discounted expected, or anticipated, returns. This rule is rejected both as a hypothesis to explain, and as a maxim to guide, investor behavior. We next consider the rule that the investor does (or should) consider expected return a desirable thing *and* variance of return an undesirable thing. This rule has many sound points, both as a maxim for, and hypothesis about, investment behavior."

19.2.2 On the surface, the logic appears impeccable. However, it should be obvious from the comments under *The Suppression of Uncertainty* in the preceding section, that 'expected return' and 'variance', as described in Markowitz (1952), are, in fact, 'perceived expected return' and 'perceived variance' on the basis of highly uncertain estimates of subjective probabilities. Should these estimates differ in any material respect from the 'true' probabilities, however defined, then deliberative action in the pursuit of so-called mean-variance efficiency may reduce, rather than enhance, the risk/reward characteristics of the portfolio. In short, mean-variance optimisation could be a recipe for what I call 'unintelligent', or type-B, behaviour, rather than for 'intelligent', or type-A, behaviour. This is precisely the point made by Grimes in the discussion on Moore (1972).

19.2.3 Another crucial point of principle is whether an investor should, indeed, always regard variance as an 'undesirable thing'. As I describe in earlier sections, higher short-term variability will lead to enhanced expected returns to investors (other than the 'unintelligent' type B) who can identify departures from 'central value' and implement dynamic strategies which exploit these pricing inefficiencies to a greater or lesser extent.

19.2.4 In the penultimate paragraph Markowitz suggests that, while tentative estimates of the future means and variances of security return might be derived from past observations, he believed that "better methods, which take into account more information, can be found." There is, however, no specific discussion of the behaviour of 'the others', and hence no theoretical framework for what I call 'optimal', or type-O, behaviour.

19.3 *Markowitz (1959)*

19.3.1 The importance of Markowitz (1959) goes far beyond its practical relevance, or otherwise, as a framework for portfolio selection, as shown, for example, by the following two testimonials on the back cover of the 1991 Second Edition:

> "With his mean-variance portfolio selection model, Harry Markowitz made the first major breakthrough in modern finance. His description and calculation of the optimum portfolio elaborated in Portfolio Selection turned out to be the indispensable building block from which the theory of the demand for risky securities was constructed."
>
> Merton H. Miller, The University of Chicago

> "Modern portfolio theory gives us a rigorous mathematical justification for the time-honored investment maxim that diversification is a sensible strategy for individuals who wish to reduce their risks."
>
> Burton G. Malkiel, Princeton University, and author of *A Random Walk Down Wall Street*

19.3.2 If we define the postulates A, B, C and D as;

A — variance of return is a satisfactory measure of risk;
B — the probability density function of return is known exactly;
C — rational behaviour is equivalent to maximising expected utility; and
D — all investors behave rationally;

and let MVO be the proposition that mean-variance optimisation provides a sound mathematical framework for portfolio selection, the case made out in Markowitz (1959) can be expressed in the following terms:

If A and B and C and D then MVO.

19.3.3 I reject postulate A for two reasons: expected return is not necessarily

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

an appropriate benchmark position from which to measure risk; and the symmetric measure of variance cannot distinguish between downside deviations (which are bad) and upside deviations (which are good).

19.3.4 I reject postulate B on the basis of my practical experience as an investment actuary. Some of the crucial limitations of a quantitative approach are described by Green in the discussion on Moore (1972):

> "Unfortunately something was always left out, be it a qualification in the auditors' report, a loss taken to profit and loss, or some equally vital but subtle piece of information. That was why a databank could never replace the report and accounts as a prime course for investment analysis, and any model which ignored the investment analyst had serious drawbacks."

19.3.5 I believe that it is this practitioners' 'sixth sense' about the incompleteness of purely numerical data that causes the 'independence axiom' to break down, thereby introducing non-linearity into the system.

19.3.6 The experience of my equities model, which is much more elaborate than most, shows that the performances of shares ranked as having similar levels of attractiveness will vary very widely, so that attempts to optimise, in a narrow statistical sense, could result in portfolio weightings that increased, rather than reduced, the general level of riskiness of the portfolio.

19.3.7 I reject postulate C on the basis that the expected utility maxim, like the existence of a 'utility curve' in the first place, is dependent on the 'independence axiom', which I regard as inconsistent with prudent financial management.

19.3.8 An interesting mathematical point is that Markowitz does not appear to assume the validity of the 'independence axiom' when proving his relatively innocuous Axioms I, II and III, which are equivalent to Axioms (3:A) and (3:B) of the von Neumann & Morgenstern system. In the last few lines of his proof of the expected utility maxim, which runs to seven pages, Markowitz does, however, use the 'independence axiom'.

19.3.9 I reject postulate D on the grounds that rationality implies equilibrium, whereas my experience of both gilts and equities has shown that very powerful cyclical elements are involved. While postulate D could lead to 'intelligent' behaviour, I believe that we should use a dynamic strategy to upgrade the potential profitability.

19.3.10 While I do not accept many of the claims made by others on his behalf, I regard Markowitz's 1959 book as an exceptionally valuable contribution, in that it attempted, for the first time, a rigorous approach to each link in the portfolio management chain. Later researchers, however, rather than examining ways in which each link could be improved, ignored much of Markowitz's painstaking discussion of the underlying assumptions that he had no alternative but to make if any formal framework was to be built at that time. In particular, he suggested that semi-variance was a more plausible risk measure than variance, he detected problems with the 'independence axiom' in lottery-type situations, and he expressed a willingness to use a better axiomatic system for rational

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

behaviour, should one become available. Also, Markowitz assumed neither stockmarket efficiency nor normal distributions in the construction of his general framework.

19.4 *Single Factor Models*

As will be apparent from my comments in Section 9, I regard the Sharpe Diagonal Model (often simply called the 'market model') as inadequate, even as an aid to logical thought, let alone as a practical tool. However, it was the combination of this single factor model and the Markowitz framework which led to the development of what is normally called Modern Portfolio Theory. It was also, at this stage, that assumptions of normality and efficiency crept in, reducing further the scope for adding value in practical applications.

19.5 *Multi-Factor Models*

While an improvement over single factor models, multi-factor models still tend, I believe, to place far too much reliance on historical data and far too little on human judgement as to the future. There is also one particular disadvantage that seems to have been overlooked. In the absence of any alternative framework for human behaviour, it is assumed that investors act intelligently, and, hence, that past price movements reflect an intelligent appraisal of available historical data. If, on the other hand, aggregate investor behaviour close to cyclical turning points in market indices or economic aggregates is, as I believe, often unintelligent, then multi-factor models, which, in essence, extrapolate past correlations, will break down completely at precisely those times when aids to investment timing could add most value. Many multi-factor models are proprietary 'black boxes', but, from diligent perusal of marketing material and numerous private conversations, I believe that this general conclusion is not unfair.

19.6 *Markowitz (1991)*

19.6.1 In 1989, at the Sessional Meeting of the Faculty of Actuaries at which Clarkson (1989) and Clarkson & Plymen (1988) were presented, Professor Markowitz commented, as follows, on his use of variance rather than semi-variance:

> "My discussion up to now has used variance as a criterion. Why do I stick to this notion? Over the telephone Mr Clarkson pointed out that in my 1959 book I spoke of something called semi-variance which was like variance but only counted the downside deviation. This is never mentioned in the 1987 book. The reason is twofold. One is, as Mr Clarkson pointed out, that it is much easier to deal with the variances of portfolios, and how they related to individual securities, than it is to use other measures of risk or dispersion ... But convenience is not sufficient. Are we coming to the right answers? ... There have been a number of experiments, cited in my 1959 book ..... The answer that has come from these experiments is that if you know mean and variance you can guess expected utility very closely."

19.6.2 After being awarded the Nobel Prize for Economics, Professor

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

Markowitz published, in 1991, a second edition of his classic 1959 book, to which were added various notes on the development of financial economics over the previous three decades or so. The 'Note on Chapter IX' contains a summary of the comments that he made at the Faculty of Actuaries Sessional Meeting in 1989, and then concludes as follows:

> "Perhaps my neglect of mean-semi-variance is not justified. Perhaps mean-semi-variance approximations are better than mean-variance approximations in certain cases where the latter leave room for perceptible improvement. Also, the *absolute* level of computer costs has fallen by orders of magnitude since 1959. The fact that mean-variance analysis is *relatively* less expensive is no longer important."

19.7 *Markowitz (1993)*

In November 1993 Professor Markowitz was the keynote speaker at a discussion meeting on 'Mathematical Models in Finance' at the Royal Society in London. At this meeting, he explained that, in recent portfolio selection work for a major securities house, he had moved to semi-variance of return against the broad stockmarket index as the appropriate measure of risk for practical portfolio selection applications. Furthermore, having realised that estimation errors in the input data could lead to unrealistically high share weightings, he incorporated maximum weightings rather than allowing these to be determined by the usual expected utility algorithms.

19.8 *Conclusions*

Although Professor Markowitz has recently found it necessary, in the pursuit of practical relevance, to abandon many of the original tenets of mean-variance optimisation, this important development seems to have escaped the attention of most financial economists. Finance theory, in general, has been unable to move beyond the highly unrealistic simplifying assumption that risk is equivalent to a statistical measure of variability of return.

20. THE EFFICIENT MARKET PARADOX

20.1 *Introduction*

20.1.1 I now address the paradox of why the financial economists of around a quarter of a century ago came to what I regard as the totally misguided conclusion that it was futile, even for professional investors, to attempt to achieve superior returns. Although some prominent financial economists, such as Fama and Rosenberg, have renounced the more extreme versions of the efficiency paradigm, modern day presentations on the subject from leading academics are still permeated with the general impression that the efficiency conclusions arrived at by earlier researchers were 'scientific'. At the Royal Society public discussion meeting in November 1993, for instance, Professor Merton of Harvard University commented, as follows, in the context of empirical studies that seemed to show

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

that professional investment managers were generally unable even to match market average returns:

> "Rigorous scientific confirmation of this belief was provided by a host of empirical performance studies along lines set by Jensen (1968) who used the CAPM as a benchmark to test for superior performance among United States mutual funds in the post war period."

20.1.2 I believe that the essential clue to the paradox lies, yet again, in the workings of the human mind, as described by Adam Smith in his *History of Astronomy*:

> "Philosophy, therefore, may be regarded as one of those arts which address themselves to the imagination; and whose theory and history, upon that account, fall properly within the circumference of our subject. Lest us endeavour to trace it, from its first origin, up to that summit of perfection to which it is at present supposed to have arrived, and to which, indeed, it has equally been supposed to have arrived in almost all former times."

20.1.3 The human mind craves certainty in an uncertain world, and thereby tends to be overconfident about the applicability of new 'scientific' approaches which, although the best available at the time, may not only be woefully inadequate as an explanation of real world behaviour, but might also be completely replaced in the near or distant future by a new and better paradigm.

20.2 *Jensen and Strong Level Efficiency*

20.2.1 The so-called evidence of 'strong level' efficiency in Jensen (1968), as cited in Merton (1993), is, perhaps, the most clear-cut example of the dichotomy between the teachings of financial economics and the conclusions that I have reached on the basis of my training and experience as both a pure mathematician and an investment actuary. Jensen's key assumptions can be summarised as follows:

- A All investors are averse to risk as measured by variability of return.
- B All investors maximise expected utility.
- C All investors have homogeneous expectations.
- D All investors use only expected return and variance of return when choosing amongst portfolios.
- E The capital market is in equilibrium.
- F Deviations from a least squares regression line of variance of return against expected return follow a normal distribution.

20.2.2 Now let Z be the 'strong level' efficiency proposition that it is futile, even for professional investors, to attempt to achieve superior returns. Then, stripping out all the highly confusing economic and statistical jargon, Jensen arrives at the following statement in terms of mathematical logic:

If A and B and C and D and E and F then Z.

20.2.3 From my training as a pure mathematician, I observe that any

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

inference of 'strong level' efficiency is invalid unless A, B, C, D, E and F are all true. From my experience as an investment actuary, I attach a very low probability, well below 0.1 in each case, to each of the assumptions A, B, C, D, E and F being true individually. Since the probability of A, B, C, D, E and F all being true is the product of the probabilities of each of these independent statements being true individually, my conclusion is that the probability of the apparent demonstration of 'strong level' efficiency in Jensen (1968) being valid is less than one in a million.

20.2.4 While Jensen's apparent demonstration of 'strong level' efficiency has less than one chance in a million of being valid in the context of my highly subjective personal probabilities, his general proposition of 'strong level' efficiency may nevertheless be true. As an 'acid text' of the competing paradigms, I now examine his detailed numerical results from the perspective of my Dynamic Equilibrium Model to see whether this exhibits greater explanatory power in terms of our understanding of real world behaviour.

20.2.5 The conceptual background of my Dynamic Equilibrium Model suggests that the aggregate risk-reward regression line obtained by Jensen will be virtually meaningless as a benchmark for investment management competence. In particular, if there are relatively few high technology or utility funds, these can be expected, for the reasons set out in ¶9.9.5, to show very bad and very good 'performances' respectively, relative to the regression line.

20.2.6 This predicted heterogeneity could not be easier to observe. Of the 115 funds investigated, the rankings of the three high technology funds were 89, 110 and 115; all below the bottom quartile, and including the very bottom fund. The one utilities fund was in first position by an implausibly high margin; its deviation from the regression line was more than two and a half times that of the fund in second place.

20.2.7 Jensen's conclusions on 'strong level' efficiency are:

> "It is important to note in examining the empirical results presented above that the mutual fund industry (as represented by these 115 funds) shows very little evidence of an ability to forecast security prices. Furthermore there is surprisingly little evidence that indicates any individual funds in the sample that might be able to forecast prices."

20.2.8 My viewpoint is that these 'conclusions' being valid only within the narrow confines of the simplifying assumptions used by Jensen, should have been accompanied by very strong warnings as to the highly tentative nature of these simplifying assumptions. No such warnings were given, with the result that the concept of 'strong level' efficiency was generally perceived as having been verified by unimpeachable scientific evidence.

20.3 *Peters, Chaos Theory, and Equilibrium*

20.3.1 I regard Peters (1991) as one of the most important books on capital market theory since Markowitz (1959), in that it brings together a vast amount of empirical evidence to show that the current paradigms of linearity, equilibrium

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

and efficiency cannot explain the very complex behaviour observed in the real financial world. Virtually all of this evidence, while anomalous in the context of the current theory of finance, is consistent with my Dynamic Equilibrium Model and the underlying hypothesis that investor behaviour can be classified into four quite distinct stereotypes.

20.3.2 Peters, as a pure mathematician, highlights the central conflict between practical relevance and the existence of analytical solutions:

> "The attempt to simplify nature by making it tidy and solvable has led to misleading conclusions. Econometric analysis was desirable because it could be solved for optimal solutions. However, if markets are nonlinear, there are many possible solutions. Trying to find a single optimal solution can be a misguided quest."

The parallels with Tolstoy's simile of the unsupervised plasterers need no elaboration.

20.3.3 Peters, as a highly experienced investment practitioner, suggests that one reason for the observed non-linearity is that people are not unbiased when they set subjective probabilities; they are likely to be more confident in their forecasts than is warranted by the information they have. My practical experience suggests that this is, indeed, a very important element of what I call type-B, or 'unintelligent', behaviour.

20.3.4 Perhaps the most powerful diagnostic tool that Peters uses is the methodology of chaos theory, and, in particular, the Hurst exponent, which provides a robust statistical measure of the time-dependency characteristics of a time series. A value significantly above 0.5 indicates a 'long-term memory' effect, which cannot be explained by any equilibrium-based theory such as that implied by the rational behaviour paradigm. The observed values of around 0.7 for both equity market indices and individual share prices, while consistent with my Dynamic Equilibrium Model, represent very strong evidence that the Capital Asset Pricing Model is a completely inappropriate theoretical framework.

20.3.5 Peters cites the very comprehensive empirical evidence in Shiller (1989), that the volatility observed in capital market series is far higher than would be expected on any equilibrium-based theory predicated on rational behaviour. Shiller suggests that this excess volatility may be due to the existence of two different types of investor, namely 'noise traders', who follow fashions and fads, and, thereby, tend to overreact to changes in fundamentals, and 'smart money traders', who invest according to value. His 'noise traders' are the equivalent of my type-B investors, and his 'smart money traders' are the equivalent of my type-A, type-O and type-R investors combined.

20.3.6 Peter concludes with the following summary:

> "We have seen evidence that the capital markets are nonlinear systems, and we have seen that current capital market theory does not take these effects into account. Because of this omission, their validity is seriously weakened. However, we do not have a full model of investment behaviour to replace the CAPM."

My categorisation of investors into four stereotypes, and the resultant theoretical construct of the Dynamic Equilibrium Model, would appear to be progress along the general lines seen as necessary by Peters if the gulf between the current equilibrium-based theory and reality is to be bridged.

20.4 *Conclusions*

I believe that the Efficient Market Hypothesis should be regarded, not as a tentative scientific conclusion about certain characteristics of the real financial world, but as a special case of the even more pervasive Equilibrium (Misguided) Hypothesis that has dominated economic science in modern times. The efficiency tests that have been carried out by financial economists are scientific only within the very narrow confines of an abstract equilibrium-based financial world that bears little resemblance to the real financial world in which I have lived and worked.

21. Adam Smith — Natural Prices and Market Prices

21.1 *Introduction*

In this section I show that the systematic departures from equilibrium that I have detected in modern day capital markets are remarkably similar to the systematic departures from equilibrium that are documented in great detail in Smith's *Wealth of Nations*.

21.2 *Impediments to Equilibrium*

21.2.1 Smith begins with a very perceptive definition of 'natural price':

> "When the price of any commodity is neither more nor less than what is sufficient to pay the rent of the land, the wages of the labour, and the profits of the stock employed in raising, preparing, and bringing it to market, according to their natural rates, the commodity is then sold for what may be called its natural price."

He then gives a very clear description of 'market price':

> "The actual price at which any commodity is commonly sold is called its market price. It may either be above, or below, or exactly the same with its natural price."

21.2.2 After discussing the general mechanisms that tend to correct any significant discrepancy between natural price and market price, Smith explains their inter-relationship as follows:

> "The natural price, therefore, is, as it were, the central price, to which the prices of all commodities are continually gravitating. Different accidents may sometimes keep them suspended a good deal above it, and sometimes force them down even somewhat below it. But whatever may be the obstacles which hinder them from settling in this centre of repose and continuance, they are constantly tending towards it."

21.2.3 Smith begins his analysis of 'wages and profit' by citing three quite distinct reasons for the very significant differences in rates that are observed — real differences in characteristics, differences in characteristics that exist only in

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

the 'imaginations of men' (i.e. systematic misconceptions), and restrictive government legislation.

21.3 *The Price of Labour*

As a result of his very extensive observations, both in Britain and on the continent of Europe, Smith concludes that there are five principal circumstances that affect wage rates — the agreeableness or disagreeableness of the employment, the difficulty and expense of learning it, its constancy or inconstancy, the degree of trust involved, and the probability or improbability of success. The most important area of systematic misconceptions that Smith identifies relates to the perceived probability of success:

> "The overwhelming conceit which the greater part of men have of their own abilities is an ancient evil remarked by the philosophers and moralists of all ages. Their absurd presumption in their own good fortune has been less taken notice of. It is, however, if possible, still more universal. There is no man living who, when in tolerable health and spirits, has not some share of it. The chance of gain is by every man more or less over-valued, and the chance of loss is by most men under-valued, and by scarce any man, who is in tolerable health and spirits, valued more than it is worth."

In the light of such unintelligent behaviour, very significant departures from economic equilibrium are likely.

21.4 *Risk Premiums for Uncertainty*

In his analysis of rates of profit in different trades, Smith observes that, while the normal rate of profit varies, as one would expect, with the certainty or uncertainty of the returns, "It does not, however, seem to rise in proportion to it, or so as to compensate it completely," and he attributes the very high rates of bankruptcies in the most hazardous trades to the same human failing of reckless overconfidence:

> "The presumptuous hope of success seems to act here as upon all other occasions, and to entice so many adventurers into those hazardous trades, that their competition reduces their profit below what is sufficient to compensate the risk."

Again, such unintelligent behaviour is likely to generate very significant departures from economic equilibrium.

21.5 *The Natural Price of Insurance*

A striking example of the similarities between actuarial thought and the principles of intelligent economic behaviour, as observed by Smith, is his conclusion that the 'natural price' of insurance consists of three elements — the net premium based on the expected claims experience, a loading for expenses, and a profit margin:

> "In order to make insurance, either from fire or sea-risk, a trade at all, the common premium must be sufficient to compensate the common losses, to pay the expense of management, and to afford such a profit as might have been drawn from an equal capital employed in any

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

common trade. The person who pays no more than this evidently pays no more than the real value of the risk, or the lowest price at which he can reasonably expect to insure it."

This invokes exactly the same logic as the decision rule that Professor Bühlmann cited in his lecture in London in June 1995 — write business if, and only if, the market premium is at least equal to the actuarial premium. The teachings of the 'individual rationality' paradigm, on the other hand, are that the natural price and the market price are always identical.

21.6 *The Natural Effort to Improve*

Although Smith was well aware of the general human failing of a propensity towards reckless overconfidence, his very extensive observations of real world behaviour led him to the conclusion that there was an even stronger positive force, namely the constant quest for self-improvement:

"The natural effort of every individual to better his own condition, when suffered to exert itself with freedom and security, is so powerful, that it is alone, and without any assistance, not only capable of carrying on the society to wealth and prosperity, but of surmounting a hundred impertinent obstructions with which the folly of human laws too often encumbers its operations."

21.7 *The* Wealth of Nations *Corn Prices*

21.7.1 Smith's pioneering approach to economic science involved six quite distinct stages — observe, classify, describe, explain, verify, and predict. In particular, he verified his tentative explanations of causal mechanisms against real world data before allowing himself to use these ideas as a framework for predicting optimal courses of future behaviour. A typical example of the verification process is in his analysis of corn prices, where, using the ten-year arithmetic average as his measure of the central price, he suggested that the fact that prices in certain recent years of oversupply had not been as low as might have been expected from an extrapolation of past experience was probably due to the effects of the 'corn bounty' which was introduced in 1688 to reduce the year-on-year variability of corn prices, and, thereby, lead to more intelligent long-term decisions as to what land and manpower resources to devote to the growing of corn. This highly perceptive analysis of deviations from a moving average must surely be the first recorded application of the principles of the mean absolute deviation approach, as described for an actuarial audience in Plymen & Prevett (1972).

21.7.2 If we take the general framework of my Dynamic Equilibrium Model, as outlined in Section 9, and assume that the strong driving force of self-improvement identified by Smith causes the proportion of unintelligent economic agents to decrease over time, we can make three predictions about the general characteristics of the annual corn prices from 1595 to 1764, as documented in Smith's *Wealth of Nations*:

(1) As measured against the geometric moving average, there will be alternate

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

runs of positive and negative deviations rather than the random distribution of positive and negative deviations implied by an equilibrium hypothesis.

(2) The typical periodic time for adjacent runs of positive and negative deviations will decrease over time.

(3) The ratio of the average deviation to the average price will show a decreasing trend over time.

21.7.3 Using the same exponential smoothing constant of 0.14 for the annual corn prices as I use for weekly equity prices, the predicted cyclical behaviour can be seen very clearly. The average periodic time decreases from just under 9 years in the first half of the period, from 1595 to 1687 (the year before the corn bounty legislation was passed), to around 5 years in the second part. Finally, the moving average of the ratio of the average deviation to the moving average of the price, while itself moving in a cyclical manner, but with a much longer typical periodic time, does, indeed, exhibit a decreasing trend over the long term.

21.8 *Conclusions*

This verification of the teachings of my Dynamic Equilibrium Model against the general characteristics of the most important economic time series of two, three and four centuries ago suggests, not only that my underlying framework for human behaviour is of fairly general validity, but also that economic science would have evolved far more rapidly, and in a far more satisfactory manner, had Smith's common sense approach of 'observe, classify, describe, explain, verify and only then predict' not been displaced by the abstract mathematical theorising that has led to unjustified presumptions of rationality and equilibrium.

22. Keynes — Probability, Confidence and Human Behaviour

22.1 *Introduction*

In this section I show that my views on how the human mind uses past experience as a guide to future behaviour are very similar to the views expressed by Keynes, both in his highly philosophical *Treatise on Probability*, published in 1921, and in his highly pragmatic *General Theory*, published in 1936.

22.2 *A Treatise on Probability*

22.2.1 Keynes (1921) stands, in my view, at the other end of the spectrum, as regards practical relevance, from von Neumann & Morgenstern (1944), as a foundation work on how past experience in general, and probability in particular, should be used as a guide to future behaviour. The former examines and attempts to disentangle all the perplexing philosophical difficulties that arise; the latter pretends that these difficulties do not exist, and, in order to arrive at numerical results, has to make the astonishingly naive assumption that all required probabilities are known exactly:

"Probability has often been visualized as a subjective concept more or less in the nature of an estimation. Since we propose to use it in constructing an individual, numerical estimation of utility, the above view of probability would not serve our purpose. The simplest procedure is, therefore, to insist upon the alternative, perfectly well founded interpretation of probability as frequency in long runs. This gives directly the necessary numerical foothold."

22.2.2 In the following paragraphs I list some of the more important observations in Keynes (1921) that I believe have to be taken into account in descriptions of real world human behaviour.

22.2.3 Jacques Bernoulli (an uncle of Nicholas Bernoulli) propounded the maxim that we should take into account all the information we have. Keynes introduces both freewill and common sense, by suggesting that it might be appropriate to search first for further relevant information before attempting to process whatever information is already available.

22.2.4 In discussing the application of probability to conduct, Keynes also explains why this exhortation to take everything into account is easily forgotten in the pursuit of numerical 'answers':

"The statistical result is so attractive in its definiteness that it leads us to forget the more vague though more important considerations which may be, in a given particular case, within our knowledge. To a stranger the probability that I shall send a letter to the post unstamped may be derived from the statistics of the Post Office; for me those figures would have but the slightest bearing upon the question."

22.2.5 The normal distribution should be used only in certain special classes of phenomena where it is known to be relevant. In particular, Pearson has shown "that the Gaussian-Laplace normal distribution is very far from being a general law of frequency distribution either for errors of observations or for the distribution of deviations from type such as occur in organic populations .... It is not even approximately correct, for example, in the distribution of barometric variations, of grades of fertility and incidence of disease."

22.2.6 In applying a least squares or similar approach, the existence of 'discordant observations' (i.e. outliers) will cause serious interpretational difficulties, and very careful consideration has to be given to the appropriateness of attaching lower (or zero) weightings to these outliers when estimating the 'best' fit.

22.2.7 In 1777 the French philosopher Buffon suggested one in ten thousand as the limit below which probability was perceived by the human mind as being negligible, on the basis that, being the probability at that time that a man of 56, taken at random, would die within a day, it would, in practical terms, be disregarded by a man of 56 who knew his health to be good. This is a highly plausible psychological rationalisation of the type of risk threshold that I describe in Section 10.

22.2.8 In the same way as a professional musician will have a far sharper sense of hearing than the vast majority of the population at large, Keynes observes that the assessment of probabilities is a human faculty where

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

achievement levels will vary markedly with training, experience and innate ability. An obvious corollary is that human behaviour will often be based on dangerously inaccurate perceptions of the true underlying probabilities.

22.3 *Classical Economic Theory*

22.3.1 My views on financial economics, at present, are identical to Keynes's views on classical economics sixty years ago, as set out in the first paragraph of his *General Theory*:

> "I shall argue that the postulates of the classical theory are applicable to a special case only and not to the general case, the situation which it assumes being a limiting point of the possible positions of equilibrium. Moreover, the characteristics of the special case assumed by the classical theory happen not to be those of the economic society in which we actually live, with the result that its teaching is misleading and disastrous if we attempt to apply it to the facts of experience."

22.3.2 My views as to how the perceived failings of the essentially linear methodologies of financial economics should be corrected are identical to Keynes's views on what required to be done, in his day, to replace the essentially linear methodologies of classical economics:

> "The classical theorists resemble Euclidean geometers in a non-Euclidean world who, discovering that in experience straight lines apparently parallel often meet, rebuke the lines for not keeping straight — as the only remedy for the unfortunate collisions which are occurring. Yet, in truth, there is no remedy except to throw over the axiom of parallels and to work out a non-Euclidean geometry. Something similar is required today in economics."

22.3.3 Keynes singled out a fundamental postulate relating to employment levels as the crucial axiom that had to be abandoned; the fundamental postulate of financial economics that I believe has to be abandoned is that of so-called 'rational behaviour'.

22.4 *The Importance of Confidence*

22.4.1 I regard Chapter 12 of Keynes's *General Theory* as by far the best account, in modern times, of human psychology, particularly as regards confidence, in the real financial world. Keynes observes that there are not, in reality, two separate factors affecting the rate of investment, namely expected return and perceived risk in terms of the 'state of confidence'; confidence effectively dominates whether investment will be contemplated or not. In other words, no matter how attractive the expected return, investment will not be considered unless the perceived probability of failure (however defined) is acceptably low. This is not only the same general situation as for physical risk in sports, but is also, I believe, the cause of the non-linearity and departures from equilibrium observed by Keynes, and others, in the real financial world.

22.4.2 In his *Treatise on Probability*, Keynes tries, but fails, to devise a numerical framework for risk and confidence. His discussion of confidence in Chapter 12 of his *General Theory* is, therefore, qualitative in nature, and, hence, impossible for mathematical economists to translate into formulae and equations:

> "There is, however, not much to be said about the state of confidence *a priori*. Our conclusions must mainly depend upon the actual observation of markets and business psychology. This is the reason why the ensuing digression is on a different level of abstraction from most of this book."

22.5 *Short-Termism and Investment Performance*

Keynes gives an extremely perceptive account of one area of financial risk involving the behaviour of others, namely the assessment of investment performance:

> "There is no clear evidence from experience that the investment policy which is socially advantageous coincides with that which is most profitable. It needs *more* intelligence to defeat the forces of time and our ignorance of the future than to beat the gun. Moreover, life is not long enough; — human nature desires quick results, there is a peculiar zest in making money quickly, and remoter gains are discounted by the average man at a very high rate."

To survive these short-term propensities of the human intellect, the investment manager may be driven by 'captains of industry' to pursue the very 'short-termism' that they abhor as seriously detrimental to long-term investment in the enterprises they manage:

> "Finally it is the long-term investor, he who most promotes the public interest, who will in practice come in for most criticism, wherever investment funds are managed by committees or boards or banks. For it is in the essence of his behaviour that he should be eccentric, unconventional and rash in the eyes of average opinion. If he is successful, that will only confirm the general belief in his rashness; and if in the short run he is unsuccessful, which is very likely, he will not receive much mercy. Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally."

22.6 *Animal Spirits*

22.6.1 Keynes summarises certain limitations of the human intellect in his usual vivid style:

> "Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as a result of animal spirits — of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities."

22.6.2 This, I suggest, is true, but incomplete. It is true in that it can be regarded as an important element of what I call type-B, or 'unintelligent', behaviour. It is incomplete in that less unintelligent future assessments will result from a combination of the strong human desire for self-improvement and a better practical framework for guiding everyday decisions in an uncertain world.

22.6.3 A very serious impediment to the acceptance of the Copernican view of astronomy was a general reluctance to believe that the earth was no longer the centre of the universe. Similarly, I suspect that Keynes's use of the phrase 'animal spirits', which appears to detract from our faith in the power of the human intellect, has resulted in his valuable insights into confidence as a driver of human behaviour receiving far less attention than they deserve.

22.7 *Conclusions*

I regard Keynes's *Treatise on Probability* and *General Theory* as the two most penetrating refutations this century of the conceptual and empirical foundations of the equilibrium-based theories that have dominated economic science in modern times.

23. The Theory of Sports

23.1 *Introduction*

23.1.1 As a result, largely, of there being no generally accepted theory of risk that can be applied to everyday practical action, human behaviour in the financial world is, in my view, so often 'unintelligent' that it would be impossible to obtain clear and uncontroversial descriptions of 'intelligent' and 'unintelligent' behaviour by direct observation of financial behaviour alone. In the physical world of potentially dangerous sports, on the other hand, where the risks involved are generally much easier to understand, the recognition and mastery of risk have evolved to such an advanced stage that the vast majority of participants, even those classed as 'beginners', follow behavioural patterns that I would describe as 'intelligent'. Sadly, however, severe injuries, or even fatalities, occur in some sports such as mountaineering and ski-ing, and, in many cases, these result from behaviour that falls short of what I would describe as 'intelligent'.

23.1.2 In this section I use further examples from sports to obtain, first of all, a description of 'intelligent' behaviour, and then, by identifying systematic shortfalls from such a standard, to pinpoint some important characteristics of 'unintelligent' behaviour.

23.2 *Intelligent Behaviour*

23.2.1 Based largely on my participation in a ski mountaineering tour in the Dolomites some years ago, I conclude that 'intelligent' behaviour in sports involves five quite distinct fields of activity. The first involves the recognition of the general characteristics of the situation, and, in the case of this tour, this involved an awareness of the skills and general level of physical fitness required and an awareness that the most serious risks were avalanches, navigational errors, which could result in being stranded high on a mountain overnight, and serious falls on unpisted slopes.

23.2.2 The second field involves preparatory action in four general areas — having sufficient technical expertise (we employed a professional guide), having the necessary specialist equipment, having adequate information about the route (such as detailed contour maps), and having an awareness of precautionary behaviour to take in high risk situations (our guide instructed us as to what action to take when crossing a slope with any meaningful avalanche risk or if actually caught in an avalanche).

23.2.3 The third field involves the identification, and then the assessment, of possible courses of action, particularly as regards whether or not, in the light of

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

the weather and snow conditions, to attempt a particular part of the route. Each plausible course of action, once identified, has to be assessed, first of all as to its expected benefits, and then as to the risk component in each of the areas of expertise, equipment, precautions and information. For example, a steep climb, normally requiring the use of ice axes, would be too risky unless all the party had experience of snow and ice climbing, and would be recklessly dangerous if the party had no ice axes.

23.2.4 The fourth field involves four quite distinct types of action. First, any instructions our guide gave had to be obeyed, as he was legally responsible for our safety on the mountains. Second, where all available courses of action involve a level of risk higher than the acceptable threshold, the one with the lowest risk is followed, regardless of the potential benefits. Third, where a life-threatening crisis, such as an avalanche, occurs, survival behaviour, in accordance with previous training, is followed. Fourth, if none of these other types applies, then 'normal behaviour', in the sense that it incorporates no specific action to avoid or minimise risk, is followed.

23.2.5 The fifth and final field, which is partly subconscious in nature, is to review past experience in order to enhance the potential for future behaviour.

23.2.6 'Intelligent' behaviour can now be summarised as follows:

Field	Component
Recognise	General characteristics
Prepare	Expertise Equipment Precautions Information
Assess	Plausible options Expected benefits Risk — expertise Risk — equipment Risk — precautions Risk — information
Act	Obey instructions Minimise high risk Survival behaviour Normal behaviour
Review	Learn from past experience

23.2.7 Since the numbers of components are 1, 4, 6, 4, 1, the binominal coefficients for $n = 4$, this formulation can be described as being of fourth order in complexity, which is very reminiscent of Keynes's comment that most investors have reached the 'third degree', and that there are some "who practice the fourth, fifth and higher degrees".

23.3 *Unintelligent Behaviour*

23.3.1 We can now define 'unintelligent' behaviour as involving any shortfall from the achievable standards of 'intelligent' behaviour that results in an unnecessarily high level of risk, and again this can be illustrated using sports examples.

23.3.2 Axiom 7 of Savage's set suggests that any event with a very low probability can be ignored, and, given the pattern recognition functioning of the human mind, this means that an event outwith previous experience and training will often be assumed to be impossible. A classic example, which clearly relates to the 'recognise' field, is that most skiers are unaware of an often significant avalanche risk.

23.3.3 In the 'prepare' field in the context of ski-ing, a lack of expertise on a steep slope can easily lead to a serious accident, not having a map and compass in very bad weather conditions can lead to an error in navigation which could prove fatal, not knowing what to do if caught in an avalanche will seriously reduce the chances of survival, and not knowing the level of avalanche risk (which is often displayed at lift stations) could result in off-piste ski-ing being attempted despite a dangerously high risk to life.

23.4 *Conclusions*

Given the pattern recognition functioning of the human mind, the formulation of 'intelligent' behaviour, sketched out above, is, I believe, a useful step forward in, first of all, understanding, and then reducing, risk. The management of risk in sports such as ski-ing is essentially a personal struggle with the elements of nature, with the behaviour of others being of only secondary importance. I see the often 'unintelligent' actions of others in the financial world, however, as introducing a significant, but generally unrecognised, element of risk, not unlike the avalanche risk in ski-ing. In the next section I shall, therefore, introduce the actions of others to derive a corresponding formulation for 'optimal' behaviour.

24. The Elements of Human Behaviour

24.1 *Introduction*

I now extend the formulation of 'intelligent' behaviour in the previous section to incorporate, first of all, the behaviour of others, and then the conscious effort to improve skill levels.

24.2 *Allowing for the Behaviour of Others*

24.2.1 In the 'prepare' category we add a further item 'know the others' as a preliminary step towards allowing in more detail for their behaviour.

24.2.2 The six items in the 'assess' category have to be re-examined, through the minds of others, to identify what choices they will analyse, and what conclusions they are likely to reach as regards expected benefits and risk components. Each of the four items in the 'act' category has to be similarly re-examined, giving a total of ten items in a new 'assess others' category.

24.2.3 In the 'assess' category, a 'predictor-corrector' approach is required. The same six items as for 'intelligent' behaviour are first of all assessed ignoring, where appropriate, the behaviour of others, and then the four items relating to the likely actions of others are assessed, giving again a total of ten items.

24.2.4 Even if the general features of the behaviour of others can be predicted, the timing will be even more difficult to predict with any accuracy. It is, therefore, necessary to add a specific additional item 'observe the behaviour of others' to the 'act' category.

24.3 *Optimal Behaviour*

24.3.1 These extensions of 'intelligent' behaviour to give 'optimal' behaviour can be summarised, as shown in Figure 24.1, where shaded circles represent the 16 items of 'intelligent' behaviour and the blank circles represent the 16 new items.

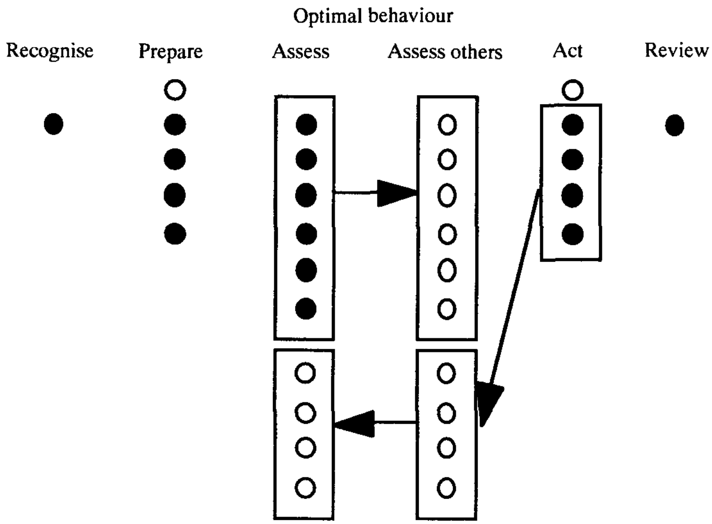


Figure 24.1.

24.3.2 We note that the numbers of items in the various categories are 1, 5, 10, 10, 5, 1, the binomial coefficients for $n = 5$. We can, therefore, describe 'optimal' behaviour as being fifth order in terms of complexity, in that it involves 2^5 items.

24.4 *Rothschild and the Battle of Waterloo*

We can now use this framework for 'optimal' behaviour to explain the Rothschild episode that I cited at the Royal Society meeting in November 1993.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

Having been able to learn of the British victory before everyone else as a result of his better 'equipment' (generally thought to have been horsemen and carrier pigeons), Rothschild reflected on how best to capitalise on his advantage. Being the George Soros of his day, he would be well aware that, if he tried to buy stock at such an acutely sensitive time in the market, others would almost certainly assume that he knew of the favourable outcome and would try to buy stock themselves, driving prices up sharply before he could deal in any significant size. If he sold, on the other hand, he reasoned that others would not double guess his strategy, but would, inferring that the British army had lost at Waterloo, trigger a stampede of panic selling. Once Rothschild judged that the height of the panic, that he did indeed create with his selling, had been reached, he bought, at far lower prices, more stock than he had earlier sold, and went out to lunch, confident in the knowledge that prices would soar as soon as official news of the victory reached the market.

24.5 *Allowing for Improvements in Skills*

In each of the 2^5 elements of 'optimal' behaviour it is clearly possible, by conscious choice, to seek to improve the level of achievement by appropriate training or experience. This conscious effort to improve, which simultaneously enhances the achievement potential and reduces the risk, upgrades behaviour from 'optimal' to 'rational'.

24.6 *Rational Behaviour*

My formulation for 'rational' behaviour is identical to that for 'optimal' behaviour, but with a 'conscious improvement' element added to each existing element. This gives a total of 2^6 elements, so that 'rational' behaviour can be described as being 6th order in terms of complexity.

24.7 *Irrational Behaviour*

We can now define 'irrational' behaviour as involving any material shortfall from the achievable standards of 'rational' behaviour.

24.8 *Conclusions*

These outline formulations of 'rational' and 'irrational' behaviour complete what I believe to be a far more powerful framework for the elements of human behaviour than that provided by current theories of finance.

25. INCOMPLETENESS PRINCIPLES

25.1 *Introduction*

In this section I develop an example in probability to derive what I see as the most fundamental statements that we are likely to be able to deduce about our understanding of the future and about rational choice under uncertainty.

25.2 Four Incompleteness Principles

25.2.1 The example in probability, which can be regarded as a formalised statement of how I believe the human mind translates past experience into a measure of confidence, is Exercise 1.50 in Gray (1967):

> "It is known that in a bag containing a large number of counters, of which some are red and some blue, there are either twice as many red counters as blue or vice versa, these two possibilities being, *a priori*, equally likely. To test which is true, a man draws a counter at random from the bag, notes the colour and repeats the process. After six trials he concludes that it is four times as likely that the bag contains the larger number of blue counters as not. Obtain the probability that, if he continues the trials, he will find his conclusion completely reversed after four more drawings, that is, that in the light of the results of all ten trials the bag is four times as likely to contain the larger number of red counters as not."

25.2.2 The required probability of $\frac{4}{81}$, or about a 5% chance, is far from insignificant. However, even if we can formulate such a precise probabilistic description of the two possible scenarios of underlying causal mechanisms, the deliberative choice that will be made as to which view will guide future action on the basis of past experience will depend on mere random chance to some degree. My first incompleteness principle is that we can never predict future human action with certainty, since it will depend on intermediate events, which themselves cannot be predicted with certainty.

25.2.3 My discussion of Donald's Paradox, in particular, and of the empirical evidence relating to probability biases in general, suggests that an adequate probabilistic description of future events is very often outwith the grasp of the untrained human mind. My second incompleteness principle is that deliberative choice under uncertainty will, as a general rule, be based on a materially incomplete understanding of the true underlying probabilities of the situation. My third incompleteness principle is that, in the absence of relevant training or practical experience over a long period of time, human action is triggered by an instinctive yes/no neural mechanism based on risk/reward assessments over which we have no conscious control. What Keynes calls 'animal spirits' is a special case of this principle.

25.2.4 Suppose now that, instead of drawing coloured counters from a bag, we base our measure of confidence on subjective probabilities relating to whether or not two or more people in different groups of 35 have the same birthday. There are 20 different groups of 35 people in hotel rooms in towns and cities around the country. If there is at least one group where no two people have the same birthday, you win a million pounds. However, there is a downside to this particular national lottery, namely that if, in every single group, at least two people have the same birthday, then your home, car and all other personal assets, worth in total one hundred thousand pounds, are sold, with the proceeds going to a national charity that you are allowed to choose. If you choose not to accept this lottery opportunity, you must participate in an alternative lottery which involves a probability of 0.9 of winning £10,000 and a probability of 0.1 of paying £1,000 to the charity of your choice, the outcome to be decided by drawing from an urn

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

containing nine white balls and one black ball. You may assume that, for a group of 35 people drawn at random from a large population where birthdays are uniformly distributed over the calendar year, the probability that at least two people have the same birthday is 0.8. Which lottery do you choose?

25.2.5 If we make the usual statistical assumptions relating to independence and the relevance of the Central Limit Theorem (generally known as the 'Law of Large Numbers'), the probability of failing to win the million pounds in the first lottery is 0.8^{20}, or about 0.01. Since this is an order of magnitude lower than the probability of failure in the second lottery, where the ratio of the possible gain to the possible loss is the same as in the first lottery, the teachings of utility theory are that we should choose the first lottery unreservedly.

25.2.6 Our instinctive 'sixth sense', however, tells us that the first lottery involves unacceptably high risk and, accordingly, that we should choose the second lottery. A few moments of deliberative thought confirm this instinctive conclusion in everyday common sense terms. First, the simplifying assumption of a uniform distribution of birthdays over the calendar year will almost certainly be wrong, but we do not know by how much. No-one playing bridge or whist, for example, would expect every hand of thirteen cards to contain precisely one ace, one king, one queen, one jack, and so on. Second, there may be other highly relevant information that is not known to the decision-maker, and various possible scenarios will come easily to mind. For example, the groups of 35 may be regular regional meetings of a national club where a condition of membership is to have been born under a particular sign of the zodiac, in which case it is certain that at least two people in each and every group of 35 will have the same birthday.

25.2.7 This 'sixth sense' paradox is my 'flagship' counter-example to Axiom 5 of Savage's suggested axiomatic framework for rational choice under uncertainty. My fourth, and undoubtedly most important, incompleteness principle is, accordingly, that the human mind, when confronted with a choice where failure could involve serious consequences, perceives an incompleteness in understanding which results in 'fail-safe' behavioural patterns that are inconsistent with the linearity implied by independence axioms of the type propounded by mathematical economists.

25.3 *Conclusions*

The general thrust of this section is that a more realistic appraisal than hitherto of the limitations of the human mind will lead to significant progress in, first of all, the understanding, and then the improvement, of human behaviour under conditions of uncertainty and risk.

26. Resolving the Robinson Crusoe Paradox

26.1 *Introduction*

Robinson Crusoe's common sense strategy of accepting the challenge if, and

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

only if, he can spot an approach which is likely to take around five minutes is equivalent to implementing the relevant elements of my fifth order framework for optimal behaviour. In particular, Crusoe recognises the likely solution times for different approaches, he quickly rejects as far too risky any approach involving a likely solution time of much more than five minutes, and he makes constructive use of the remainder of the available time to see if he can spot an approach that will allow him to exploit the opportunity with an acceptably low probability of failure. He has also prepared himself mentally for the scenario where he accepts the challenge, but has not completely solved the puzzle as the 15 minutes deadline approaches. There is a subtle point here regarding the behaviour of the others. The natives will accept the correct answer as a convincing demonstration of his intellectual powers, even if it was a pure guess. There are obvious parallels in the real world when investment performance is being assessed.

26.2 *Utility Theory*

Since the levels of risk differ enormously, depending on whether or not the challenge is accepted, utility theory is of no relevance whatsoever in this choice under uncertainty.

26.3 *The Stochastic Model*

The reason for the stochastic model 'solution' differing from the common sense strategy can now be seen very easily. The scenarios involving expected solution times of 10, 15 and 20 minutes are discarded very quickly as plausible courses of action, on grounds of involving unacceptably high risk, but these scenarios contribute virtually all of the total probability of failure of around 0.4 that is implied by the stochastic model. Basing future choice on this probabilistic and fatalistic 'all worlds' description, obtained from observations of the past, assumes that precisely what has happened in the past will happen in the future and, thereby, suppresses completely the freewill of the individual to abstain from courses of action that are perceived to involve unacceptably high risk. As with utility theory, the crucial oversight, in what might appear to some to be a scientific probabilistic approach, is that it fails to recognise the vital element of freewill encapsulated in my incompleteness principles.

26.4 *Conclusions*

My Robinson Crusoe example is, I believe, not untypical of many aspects of choice under uncertainty in the real financial world. I therefore conclude that utility theory and stochastic modelling (other than in the analysis of specific scenarios) should not be used in the construction of a scientific framework for choice under uncertainty and risk in the financial world. My new framework for identifying, and then implementing, the elements of rational behaviour is, I believe, a better way forward.

27. RATIONAL FINANCIAL BEHAVIOUR

27.1 *Introduction*

To illustrate how my framework for rational behaviour can be used to reduce the likelihood of serious financial accidents, I shall consider its application to trading in derivatives against a benchmark of cash.

27.2 *The New Approach*

27.2.1 The crucial first step is to consider the general characteristics of the situation. In the very broadest of terms, derivatives can be used either to take a long or short position in a market or to exploit perceived arbitrage opportunities between two similar financial or economic entities. In both cases, a very high exposure to both potential profits and potential losses can be achieved for a relatively small outlay in terms of initial margin. However, should values move adversely, further capital has to be provided by way of variation margin on a daily basis.

27.2.2 Where either long or short positions in a market are taken, it must be recognised that, over the short and medium terms, markets can move to extreme levels of overvaluation or undervaluation which, in hindsight, are regarded as having been grossly inconsistent with the underlying fundamentals. Equity market levels at December 1974 and October 1987 are classic examples.

27.2.3 Where arbitrage positions are taken, a clear distinction must be made between past behaviour and likely future behaviour. A classic example is the arbitrage switching that took place around a quarter of a century ago between British Petroleum and Burmah, when Burmah had a very large holding of British Petroleum shares and relatively insignificant profits from other sources. Burmah then developed, on a highly geared basis, other trading activities which generated sufficient profits to take the share price well above the previous arbitrage basis, until a concatenation of highly adverse circumstances led to the company having to be rescued by the Bank of England on 31 December 1974. In more recent times, similar salutory examples have involved previously stable relationships between bond markets or currencies in different countries breaking down as a result of divergences in particular aspects of economic performance. The exit of sterling from the Exchange Rate Mechanism in September 1992 was a particularly vivid example of this effect.

27.2.4 As regards expertise, it will, in general, be the most senior and longest serving officers in a financial institution who can best assess the likelihood of the breakdown of previously stable and potentially profitable arbitrage opportunities. It is essential that they should not only protect the financial survival of the institution by specifying maximum exposure limits based on a pragmatic assessment of historic worst case experience, but should also actively discourage any large trading positions which they believe, from their perspective of longer practical experience, to be misguided.

27.3 *Current Theory*

27.3.1 Current financial theory, on the other hand, tends to measure risk in a purely quantitative manner on the basis of a statistical portrayal of past behaviour which cannot accommodate either random shocks or the breakdown of previously stable differentials. Mandelbrot (1977) describes the unsoundness of this approach in the following terms:

> "Faced with a statistical test that rejects the Brownian hypothesis that price changes are Gaussian, the economist can try one modification after another until the test is fooled. A popular fix is censorship, hypocritically labeled 'rejection of statistical outliers'. One distinguishes the ordinary 'small' price changes from the large changes that defeat Alexander's filters. The former are viewed as random and Guassian, and treasures of ingenuity are devoted to them ..., as if anyone cared. The latter are handled separately as 'nonstochastic'."

27.3.2 Mandelbrot then states, as an alternative approach, his scaling principle, which is fully consistent with my Dynamic Equilibrium Model, but totally inconsistent with the standard assumptions of financial economists:

> "At the opposite of the fixes stands my own work. It applies to diverse data of economics, but the principle is best expressed in the context of price.
>
> SCALING PRINCIPLE OF PRICE CHANGE
>
> When $X(t)$ is a price, $\log X(t)$ has the property that its increment over an arbitrary time lag d, $\log X(t + d) - \log X(t)$, has a distribution independent of d, except for a scale factor."

27.3.3 If, as seems often to be the case, trustees, directors and others in senior positions are unaware of the precise nature, and susceptibility to unforeseen shocks, of the positions taken by traders, derivatives trading against a benchmark of cash, using standard financial models to control risk can be likened to an inverse lottery, where, in general, a series of moderate profits may, from time to time, be followed by a massive loss when previously stable differentials suddenly break down for reasons that a trader could not have been expected to foresee.

27.4 *Conclusions*

Given its explicit recognition that thresholds of downside risk should not only be set, but also monitored, by those within an organisation who have the most extensive practical experience, I believe that my framework for rational behaviour, rather than the current theory of finance, offers the better basis for the establishment of appropriate prudential controls in high risk areas such as derivatives trading.

28. Investment Management

28.1 *Improving the Performance*

28.1.1 Whether or not a detailed equity selection model of the type described

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

in Section 8 is used, various pointers for the management of equity portfolios can be deduced from my formulation of rational behaviour, especially in the 'prepare' field areas of expertise and information, and in the 'assess others' field.

28.1.2 As regards expertise, two very general observations can be made. First, the potential for successful decisions at the stock selection level will increase, both with the number of investment analysts involved and with their skill and experience. Second, investment analysts, while highly knowledgeable about the sectors and companies they research, may have limited experience of investor psychology and, in particular, of how an apparently undervalued share can remain undervalued for some considerable time before enough other investors appreciate its merits. Rather than construct portfolios using analysts' expected returns taken at face value, the assessment of the likely behaviour of other investors has to be incorporated into the investment process. This is, indeed, what generally happens in practice, and most large investing institutions have a structured decision-making process comprising analysts, portfolio managers, head of equities and head of investment, virtually identical to the army structure of privates, junior officers, generals and commander-in-chief described by Tolstoy.

28.1.3 In terms of external information, the more that is available in terms of economic reports, stockbrokers' research on sectors and companies, and current market prices and relative ratings, the less uncertain should be the forecasts on which investment strategy is based, and the greater should be the scope for superior performance. Two highly desirable areas of internal information are accurate statements, available at different levels of detail, of the current structure of portfolios, and reliable analyses of portfolio performance down to sector and individual stock levels.

28.1.4 The 'assess others' field is of vast practical importance, and requires, first of all, the identification, and then the analysis, of plausible near-term scenarios for the collective action of both institutional and private investors.

28.2 *Assessing the Performance*

28.2.1 In the absence of any recognition by financial economists, or other academics, that short-term market movements may be inconsistent with the longer-term fundamentals, trustees and others with responsibility for investment performance generally assume that relative performance over each three months period is a valid indicator of the competence of the investment manager. However, this strong emphasis on short-term performance can often hinder, rather than help, the quest for superior long-term performance, by favouring the 'unintelligent' as opposed to the 'intelligent' investment manager. If, as will often be the case, the 'intelligent' manager has a poor quarter, so much time and effort will be spent explaining the apparently unsatisfactory situation to clients that far less time can be spent on investment management, to the likely detriment of the future performance. When the short-term performance of the 'intelligent' investment manager is so poor that the investment contract is terminated, perhaps just before a very sharp market movement which fully vindicates the recent

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

stance, the funds generally move to a manager pursuing a diametrically opposed strategy which, although recently successful, may soon revert into serious underperformance. There is the further loss to clients in the aggregate of significant transaction costs associated with the realignment of the portfolios.

28.2.2 The potential disutility to society as a whole is then compounded by the 'risk equals variability of return' paradigm, which translates the higher volatility caused by 'unintelligent' participants into a perception of higher risk, with the result that less available capital is devoted to new wealth-creating equity investment.

29. Asset Liability Modelling

29.1 *Introduction*

Since asset-liability modelling, which covers both the management and the regulation of life assurance companies, general insurance companies and pension funds, is so vast and diverse an area, I restrict myself in this section to a few very general comments.

29.2 *The New Approach*

29.2.1 The approach outlined in Section 16 stands or falls depending on whether or not stable central values of market yields can be identified. In early 1995 I investigated U.K. market data and concluded that $3\frac{3}{4}\%$ for the real yield on index-linked gilts and $4\frac{1}{4}\%$ for the dividend yield on U.K. equities represented reasonably stable central values. Shortly thereafter it was announced that precisely the same values were to be used as the reinvestment yields for projections of the Minimum Funding Requirement in terms of the Pensions Act (1995).

29.2.2 It is interesting to note that the simple model for U.K. pension funds, used in Loades (1995), defines risk in a downside manner as the probability of the solvency ratio falling below a specified value, taken as 100% or 90%.

29.3 *Stochastic Models*

29.3.1 In view of the acute non-linearity introduced by the actions of 'unintelligent' participants, I regard linear models of the Box-Jenkins or ARIMA type as highly misleading and potentially dangerous representations of real world financial behaviour. A probabilistic map of deviations from central values of market yields, based in the first instance on historic data and then smoothed and adjusted to incorporate likely future trends, seems to be not only a far more practical approach, but also one which is fully consistent with the '*certum ex incertis*' and '*ad finem fidelis*' traditions of the U.K. actuarial profession.

29.3.2 Of far more practical relevance, I believe, is my conjecture, for the reasons set out in Section 26, that stochastic models should be used only to analyse specific scenarios. Their attempted application to probabilistic mixtures of scenarios, where probability distributions are essentially a reflection of past

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

history, not only overrides human judgement as to likely future behaviour, but also generates spuriously high values of variability of return that are translated into admonitions in favour of 'low risk' behaviour that reduces significantly the likely long-term returns.

29.4 *Conclusions*

The Pensions Act (1995) presents significant opportunities to the U.K. actuarial profession. The corresponding threat is that the linear methodologies of financial economics, and in particular the equilibrium-based symmetric approach to risk, may fall far short of the approaches that, in due course, are perceived to have been the most appropriate in this crucial application area of actuarial science.

30. Capital Projects

Although the diversity of characteristics of individual projects will be very great, my formulations of optimal and rational behaviour offer structured frameworks for human judgement. The 'recognise' field includes the identification of the differing objectives and risk tolerances of the various parties. In the case of Private Finance Initiative projects, for instance, the objectives of the relevant government bodies will be essentially social in nature, with risk perceived as failure to meet functionality or standards of service, whereas the objectives of the main contractor will relate to return on capital employed, with risk perceived as a shortfall from a target rate of return. The 'prepare' field includes the identification and recruitment of the necessary expertise and identifying and obtaining available information relating to previous projects of a similar nature. The 'assess' field includes the identification and analysis of various possible scenarios, and the 'assess others' field introduces the dynamic interplays of factors such as likely competition. The first crucial item in the 'act' field is the yes/no decision on viability, depending on whether or not the objectives of all the parties can be met within acceptable levels of perceived risk.

31. A Periodic Table for Economic Science

31.1 *Degrees of Uncertainty*

31.1.1 The first step towards completing a periodic table of the general type described in ¶6.9.3, is to specify a scale for areas of economic and financial behaviour in terms of increasing degrees of uncertainty and hence of complexity. The first three areas of behaviour can be identified fairly easily:

Degree of uncertainty	Area of behaviour
I	Fixed interest prices in relative terms
II	Equity prices in relative terms
III	Fixed interest and equity market levels

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

31.1.2 For the next two areas I use 'economic behaviour' and 'economic management' respectively. The former covers both 'economic choice' by an individual or institution and the observed behaviour of economic variables (such as inflation), while the latter covers the action of governments, central banks, regulators, and other supervisory authorities in their efforts to control economic and financial behaviour for the general good of society as a whole.

31.1.3 For the final area I use 'choice of economic system', since I perceive this to be the area of behaviour where the most serious conflicts can arise between long-term benefit to society as a whole and the short-term interests of those who control the choice. Adam Smith highlights this type of conflict in a passage near the beginning of his *Lectures on Jurisprudence*:

> "It is the sense of public utility, more than of private, which influences men to obedience. It may sometimes be for my interest to disobey, and to wish government overturned, but I am sensible that other men are of a different opinion from me, and would not assist me in the enterprise. I therefore submit to its decision for the good of the whole."

31.1.4 The columns of the periodic table can now be described as follows:

Degree of uncertainty	Area of behaviour
I	Fixed interest securities (relative)
II	Variable income securities (relative)
III	Capital market levels
IV	Economic behaviour
V	Economic management
VI	Choice of economic system

31.2 *Classification*

31.2.1 In terms of level of understanding, I classify utility theory as Level 1 for two main reasons, the unsatisfactory nature of some of the key axioms, and the absence of any explicit allowance for downside risk.

31.2.2 I classify MPT as Level 1 for three reasons. It incorporates a very primitive approach to risk, namely the variability of return with no distinction between the downside and the upside, it has a simplistic single period time horizon which will not necessarily be anywhere near optimal, and high share weightings obtained in the 'optimisation' process will often be no more than a reflection of estimation errors in the input data. However, since the practical approach, described by Markowitz in 1993, addresses each of these three aspects, I classify this specific approach, but not the generic mean-variance framework set out in Markowitz (1959), as Level 4.

31.2.3 Because of its success (as with Kepler's laws of planetary motion) in describing real world behaviour in mathematic terms, I classify the empirical Black-Scholes option pricing formula (but not its theoretical justification in terms of replicating strategies or 'risk free world' arguments) as Level 4.

31.2.4 I classify all other approaches within mainstream financial economics as no higher than Level 2, for five general reasons, not all of which may be relevant in any particular case. First, the frameworks used for risk do not

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

incorporate a downside approach which takes explicit account of the circumstances of the individual or institution. Second, there is no recognition of differing levels of behaviour as a result of differing levels of skill and experience. Third, there is a general dependence on utility theory, which I place at Level 1. Fourth, the incorporation of a limited degree of non-linearity through ARCH, GARCH and similar methodologies is not, I believe, anywhere near radical enough to reflect the acute non-linearity often observed in the real world. Fifth, there is no general recognition of non-stationarity in historic statistical data, let alone any attempt to identify likely causes as a guide to possible future changes in the systems.

31.2.5 I classify the Wilkie model as Level 2. By decomposing equity prices into a dividend series and dividend yields, it is much more powerful than pure diffusion models, which I would classify as Level 1. However, for the reasons set out in Section 12, I do not believe that the inflation model, even with a modification to include ARCH effects, is a sufficiently realistic description of real world behaviour to justify inclusion at Level 3. This is consistent with comments in the final section of Wilkie (1995), to the effect that areas of further work could include investigations of why observed equity market cyclicality is higher than predicted and of the psychological factors which drive capital market levels. In view of these comments, I also include Wilkie at Level 3, which relates to the recognition that a new paradigm is required if we are to achieve a satisfactory level of understanding of real world behaviour.

31.2.6 Also at Level 3 I include Mandelbrot for his demonstrations that financial and economic series do not behave in line with the standard assumptions of financial economists, I include Shiller for his work on 'excess volatility', and I include Tversky for his questioning of the assumptions underlying rational behaviour. I also include Mandelbrot at Level 4 for his scaling principle, which is not dissimilar, in broad conceptual terms, to Kepler's first law.

31.2.7 I include Allais at Level 3 for his strong criticism of the axiomatic approach to utility theory and at Levels 4 and 5 for his very detailed analysis of human behaviour under conditions of uncertainty and risk and for the importance that he attaches to psychological driving forces.

31.2.8 I include Keynes at Degrees III, IV and V of Level 5 for his very detailed analyses of human behaviour in the areas of stockmarket investment, business management and economic management by governments.

31.2.9 I include Smith at Degrees IV, V and VI of Level 5 for his very detailed analyses of human behaviour in all areas of economic affairs.

31.2.10 Although his work is not discussed earlier in the paper, I include Lucas, who was awarded the Nobel Prize for Economics in October 1995, at Degree V of Level 3. One of the main themes running through his work is that individuals, by changing their behaviour in response to actual or anticipated government action, will often nullify, to a greater or lesser extent, the perceived benefits of that economic policy. This theme is a very important special case of the complex interactions that result from my differing levels of behaviour;

Degree of uncertainty

Level of understanding	I	II	III	IV	V	VI
1	UT	UT, MPT	UT			
2	FE	FE	FE, W	FE, W		
3			Mb, Sh, W	A, Mb, T	L	
4	P	WH, BS, M	Mb	A, Mb		
5			K	S, K, A	S, K	S
6						

A	Allais	P	Pepper
BS	Black-Scholes formula	S	Smith
FE	Financial economics	Sh	Shiller
K	Keynes	T	Tversky
L	Lucas	UT	Utility theory
M	Markowitz–1993	W	Wilkie
Mb	Mandelbrot	WH	Weaver & Hall
MPT	Modern Portfolio Theory		

Figure 31.1.

'optimal' and 'rational' behaviour on the part of some individuals can defeat the objectives of 'intelligent' behaviour on the part of a government.

31.2.11 Finally, to incorporate some further actuarial reference points in addition to the work of Wilkie, I include Pepper and Weaver & Hall at Level 4 for their very detailed gilt-edged and equity models respectively.

31.3 *Relative Power*

31.3.1 There are very strong parallels between the metallic or non-metallic nature of a chemical element and the usefulness or otherwise of an economic theory. The essential characteristic of a metal is the capability to lose its free electrons to form a stable bond with other compatible elements. In the periodic table of chemical elements, this capability decreases from left to right along rows as electron shells fill, and increases down columns as atomic number increases. The essential characteristic of a useful economic theory is the capability of individuals or organisations to implement appropriate strategies for economic choice under conditions of risk and uncertainty. In the periodic table of economic theories, this capability decreases from left to right along rows as the degree of uncertainty increases and increases down columns as the level of understanding increases.

31.3.2 We can use these parallels to define a ranking measure of the usefulness of an economic theory in its specific area of application as its level of

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

understanding less the degree of uncertainty in that area. I interpret a 'relative power' greater than zero as 'very useful', a 'relative power' of zero, –1 or –2 as 'moderately useful', and a 'relative power' of –3 or lower as an indication that no useful theory exists. I exclude from my periodic table approaches that fall into the 'no worthwhile theory' category.

31.3.3 Since I regard the theory of games as a special case of utility theory at Degree V, it comes into my 'no worthwhile theory' category, and hence is excluded from my periodic table.

31.4 *The Periodic Table*

My periodic table of economic theories, with 'moderately useful' theories in the three central diagonal rows and 'very useful' theories in the bottom left area, is set out in Figure 31.1, together with a key to the individual elements.

32. CONCLUSIONS

32.1 *Differing Viewpoints*

32.1.1 The longer-term benefits to society as a whole of developing an approach which, in the words of Keynes, can "defeat the dark forces of time and ignorance which envelop our future" have to be weighed against the shorter-term benefits to an individual or academic institution of being able to construct a mathematical model from simplifying assumptions that are taken at face value. My personal viewpoint, based on a firmly held belief that the prevalence of 'unintelligent' behaviour increases very rapidly with the degree of uncertainty, is that much more weight should be given to potential utility than to mathematical tractability.

32.1.2 Mathematicians and economists, working essentially in an academic environment, will have little or no reason to believe that the standard assumptions of the rational behaviour and equilibrium paradigms may be inconsistent with observed real world behaviour, and they can, therefore, be expected to have a quite different viewpoint, namely that research resources should be directed to approaches where the scope for mathematical development is high rather than to following the pioneering work of Smith, Keynes and Allais in grappling with the conceptual difficulties posed by departures from rational behaviour and equilibrium. Papers by financial economists very rarely make reference to the works of these three individuals.

32.1.3 Of the 13 papers presented at the Royal Society discussion meeting in November 1993, only one, a philosophic essay by Harding entitled 'Making Money from Mathematical Models', makes any reference to the work of Smith. After describing how Smith's approach was "deeply empirical, founded on detailed and analytical observation of the nature of human society", Harding draws a parallel with Newton's work in the physical sciences:

> "This model, its laws of supply and demand and its concept of the division of labour have played the role slightly akin to that Newton's laws of motion played in physics."

These comments tie in very well with my classification of Smith's work as Level 5, corresponding to Newton's discovery of the causal mechanism of gravity as the unifying principle in astronomy.

32.2 *Levels of Behaviour*

32.2.1 The first unifying principle that I believe should be incorporated within economic science is the recognition, not only of differing levels of behaviour, but also of the innate ability of an individual or institution, first of all to eliminate highly disadvantageous elements of 'unintelligent' behaviour and thereafter to move on to 'optimal' and then 'rational' behaviour.

32.2.2 This recognition within economic science of differing levels of behaviour would have strong parallels with a long-established fundamental principle of actuarial science, namely the recognition that mortality data have to be classified into distinct sub-groups before further analysis if there are any grounds for suspecting that heterogeneity might exist. An aggregate statistical analysis of an amalgam of assured life, annuitant and population mortality data would be unsound in the extreme, and could have no practical value. However, in the absence of any alternative and more detailed models of human behaviour, most of the investigations carried out within the current frameworks of economic science in general, and financial economics in particular, are of a similar nature, in that they assume — despite overwhelming empirical evidence to the contrary — that all economic agents act rationally.

32.3 *Risk and Probability*

32.3.1 The second unifying principle that I believe should be incorporated within economic science is a 'downside' approach to risk, which not only incorporates an appropriate maximum risk threshold, but also takes explicit account of the possibility of 'unintelligent' behaviour on the part of others. I am not aware of any methodologies within economic science in general, or financial economics in particular, which come anywhere near meeting what I see as these crucial requirements.

32.3.2 An inevitable consequence of the complexity of real world human behaviour is that any attempts to apply this new downside theory of risk in precise probabilistic terms will be doomed to failure. There are too many degrees of freedom and too many causes of non-stationarity for a purely analytic solution, based on probability distributions, to have any meaningful practical relevance. A combination of pattern recognition, ranking measures based mainly on past experience and common sense rules is likely to be a far better way forward.

32.4 *Concluding Remarks*

32.4.1 The unmistakable conclusion that I draw from the periodic table built up in the previous section is that the current theories of financial economics are seriously incomplete as regards a scientific framework that could be used to guide prudent financial behaviour. I see three crucial areas of incompleteness — no general recognition that financial behaviour in the real world is far more complex than that implied by the standard simplifying assumptions, no recognition of differing levels of behaviour as a result of differing levels of skill and experience and the absence of a credible theory of risk.

32.4.2 This paper is, of course, also seriously incomplete. Just as Rome was not built in a day, a comprehensive new theory of finance cannot be constructed overnight. Accordingly, my main objective in writing this paper has been to show that a much better theory, and, moreover, one that is essentially actuarial in nature, is well within our grasp.

32.4.3 My overall conclusion is that the general teachings and methodologies of financial economics are dangerously unsound in principle, and thereby fall far short of the very high standards of prudent financial management required for the future success of the U.K. actuarial profession.

REFERENCES

ALLAIS, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Americaine. *Econometrica*, **21**, 503-46.

ALLAIS, M. (1954). L'outil mathématique en economique. *Econometrica*, **22**, 58-71.

ANAND, P. (1993). *Foundations of rational choice under risk.* Clarendon Press, Oxford.

ARROW, K.J. & DEBREU, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica*, **22**, 265–290.

BERNOULLI, D. (1738). Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanea*, **5**, 175–92. English translation in *Econometrica*, **22** (1954), 23–36.

BERNOULLI, D. (1760). Essai d'une nouvelle analyse de la mortalité cause pur la petite vérole. *Hist. de l'Acad., Paris*, 1–45.

BLACK, F. & SCHOLES, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, **81**, 637–654.

BOWERS, L.N., GERBER, H.U., HICKMAN, J.C., JONES, A.J. & NESBITT, C.J. (1986). *Actuarial mathematics.* The Society of Actuaries.

BÜHLMANN, H. (1987). Actuaries of the third kind? *ASTIN Bulletin* 17.

BURMAN, J.P. & WHITE, W.R. (1972). Yield curves for gilt-edged stocks. *Bank of England Quarterly Bulletin*, December 1972.

CLARKSON, R.S. (1978). A mathematical model for the gilt-edged market. *T.F.A.* **36**, 85–160 and *J.I.A.* **106**, 85–148.

CLARKSON, R.S. (1981). A market equilibrium model for the management of ordinary share portfolios. *T.F.A.* **37**, 349–607 and *J.I.A.* **110**, 17–134.

CLARKSON, R.S. (1989). The measurement of investment risk. *T.F.A.* **41**, 677–750 and *J.I.A.* **116**, 127–178.

CLARKSON, R.S. (1990). The assessment of financial risk. *Transactions of the 1st AFIR International Colloquium*, Paris, **2**, 171–194

CLARKSON, R.S. (1991). A non-linear stochastic model for inflation. *Transactions of the 2nd AFIR International Colloquium*, Brighton, **3**, 233–253.

CLARKSON, R.S. (1993). Some observations on the theory of games. *Transactions of the 3rd AFIR International Colloquium*, Rome, **2**, 507–17.

CLARKSON, R.S. (1994). The coming revolution in the theory of finance. *Transactions of the 4th AFIR International Colloquium*, Orlando, **1**, 1–13.

CLARKSON, R.S. (1995a). Some observations on the Black-Scholes option pricing formula. *Transactions of the 5th AFIR International Colloquium*, Brussels, **3**, 1125–1136.

CLARKSON, R.S. (1995b). Asset/liability modelling and the downside approach to risk. *Transactions of the 25th International Congress of Actuaries*, Brussels, **3**, 139–163.

CLARKSON, R.S. & PLYMEN, J. (1988). Improving the performance of equity portfolios. *J.I.A.* **115**, 631–691 and *T.F.A.* **41**, 631-675.

COOTNER, P.H. (1962). Stock prices: random vs. systematic changes. *Industrial Management Review*, **3**, 24-45.

DONALD, D.W.A. (1970). *Compound interest and annuities certain.* 2nd edition, Heinemann, London.

EINSTEIN, A. (1920). *Relativity.* Routledge, London.

ENGLE, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates at the variance of United Kingdom inflation. *Econometrica*, **55**, 987-1007.

FAMA, E.F. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, **25**, 383–417.

GEOGHEGAN, T.J., CLARKSON, R.S., FELDMAN, K.S., GREEN, S.J., KITTS, A., LAVECKY, J.P., ROSS, F.J.M., SMITH, W.J., TOUTOUCHINI, A. (1992). Report on the Wilkie investment model. *J.I.A.***119**, 173–228.

GRAHAM, B. & DODD, D. (1951). *Security analysis.* McGraw–Hill, New York.

GRAY, J.R. (1967). *Probability.* Actuarial Tuition Service, London.

HARDING, D. (1993). Making money from mathematical models. *Philosophic Transactions of the Royal Society of London*, Series A, **347**, 511-15.

HEILBRONER, R.L. (1986). *The essential Adam Smith.* Oxford University Press.

HEMPSTED, J.R. (1961). The expected yield of ordinary shares, *J.S.S.* **16**, 401–426.

HUGHES, D.W. (1987). The history of Halley's Comet. *Philosophic Transactions of the Royal Society of London*, Series A, **323**, 349-367.

HODGES, S.D. & BREALEY, R.A. (1971) Portfolio selection in a dynamic and uncertain world. Seminar on Analysis of Security Prices, May 1971, Chicago.

JARROW, R.A. & RUDD, A. (1983). *Option pricing.* Irwin.

JENSEN, M.C. (1968). The performance of mutual funds in the period 1945-64. *Journal of Finance*, **23**, 389–416.

KEETON, W.T. (1972). *Biological science.* Norton, New York

KEYNES, J.M. (1921). *A treatise on probability.* Macmillan, London.

KEYNES, J.M. (1936). *A general theory of employment, interest and money.* Macmillan, London.

KUHN, T.S. (1970). *The structure of scientific revolutions.* The University of Chicago Press.

LOADES, D.H. (1995). Mismatching pensioner liabilities–surplus and solvency. *Transactions of the 25th International Congress of Actuaries,* Brussels, **3**, 527-550.

MACHINA, M.J. (1987). Choice under uncertainty: problems solved and unsolved. *Journal of Economic Perspectives*, **1**, 121–54.

MANDELBROT, B.B. (1963). The variation of certain speculative prices. *Journal of Business*, **36**, 394–419.

MANDELBROT, B.B. (1977). *The fractal geometry of nature.* Freeman, New York.

MARKOWITZ, H.M. (1952). Portfolio selection. *Journal of Finance*, **7**, 77–91.

MARKOWITZ, H.M. (1959). *Portfolio selection.* Second edition (1991), Blackwell, Oxford.

MARSHALL, J.B. (1953). British government securities. *T.F.A.* **22**, 19–96.

MERTON, R.C. (1993). Influence of mathematical models in finance on practice: present and future. *Philosophic Transactions of the Royal Society of London*, Series A, **347**, 451–63.

MILLER, M.H. & MODIGLIANI, F. (1961). Dividend policy, growth and the valuation of shares. *Journal of Business*, **34**, 411–33.

MILLS, T.C. (1991). Equity prices, dividends and gilt yields in the U.K.: cointegration, error correction and 'confidence'. *Scottish Journal of Political Economy*, **38**, 242–55.

MOORE, P.G. (1972). Mathematical models in portfolio selection. *J.I.A.* **98**, 103–148.

von Neumann, J.R. & Morgenstern, O. (1944). *Theory of games and economic behavior.* Princeton University Press.

Osborne, M.F.M. (1959). Brownian motion in the stock market. *Operations Research*, **7**, 145–173.

Pegler, J.B.H. (1948). The actuarial principles of investment. *J.I.A.* **74**, 179–211.

Penrose, R. (1989). *The emperor's new mind: concerning computers, minds, and the laws of physics.* Oxford University Press.

Pepper, G.T. (1964). The selection and maintenance of a gilt-edged portfolio. *J.I.A.* **90**, 63–103.

Pepper, G.T. (1994). *Money, credit and asset prices.* Macmillan, London.

Peters, E.E. (1991). *Chaos and order in the capital markets.* Wiley, New York.

Plymen, J. & Prevett, R.M. (1972). The computer for investment research. *T.F.A.* **33**, 143–186.

Ronan, C.A. (1981). *The practical astronomer.* Roxby Press, London.

Rothschild, M. & Stiglitz, J. (1970). Increasing risk: I. a definition. *Journal of Economic Theory*, **2**, 225–43.

Rothschild, M. & Stiglitz, J. (1971). Increasing risk: II. its economic consequences. *Journal of Economic Theory*, **3**, 66–84.

Rudd, A. & Clasing, H.K. (1982). *Modern portfolio theory.* Dow Jones–Irwin, Homewood, Illinois.

Royal Society of London (1992). Risk: analysis, perception and management. Report of a Royal Society Study Group.

Savage, L.J. (1954). *Foundations of statistics.* Wiley, New York.

Schrödinger, E. (1967). *What is life?* Cambridge University Press.

Sharpe, W.F. (1970). *Portfolio selection and capital markets.* McGraw–Hill, New York.

Shiller, R.J. (1989). *Market volatility.* Massachusetts Institute of Technology.

Smith, A. (1978). *An enquiry into the nature and causes of the wealth of nations. Volume II of the Glasgow Edition of the Works and Correspondence of Adam Smith.* Oxford University Press.

Smith, A. (1978). *Essays on Philosophical Subjects. Volume III of the Glasgow Edition of the Works and Correspondence of Adam Smith.* Oxford University Press.

Tolstoy, L. (1943). *War and peace.* Translated by L. and A. Maude, Macmillan, London.

Tversky, A. (1990). *The psychology of risk. In quantifying the market risk premium phenomena for investment decision making.* Institute of Chartered Financial Analysts, Charlottesville, VA.

Weaver, D. & Hall, M.G. (1967). The evaluation of ordinary shares using a computer. *J.I.A.* **93**, 165–227.

Wilkie, A.D. (1984). Steps towards a comprehensive stochastic investment model. *O.A.R.D.P.* 36.

Wilkie, A.D. (1986). A stochastic investment model for actuarial use. *T.F.A.* **39**, 341–403.

Wilkie, A.D. (1995). More on a stochastic investment model for actuarial use. *B.A.J.* **1**, 777–977.

Williams, J.B. (1938). *The theory of investment value.* Harvard University Press.

ABSTRACT OF THE DISCUSSION

The President (Mr G. M. Murray, C.B.E., F.F.A.): welcomed visitors, including an official guest of the Faculty, Mr Barry Riley.

Professor R. S. Clarkson, F.F.A. (introducing the paper): It is both a very great privilege and a very great honour for me to be given the opportunity tonight to contribute to a crucial debate: which mathematical frameworks should actuaries use to model capital market behaviour and financial risk? Most authors of sessional meetings papers can take for granted the basic principles of their subject, and then concentrate on applying these generally accepted principles to practical problems in the real world. My task, unfortunately, is the much more difficult one of trying to offer constructive criticism on the underlying principles of methodologies which have already achieved overwhelming general acceptance in many quarters, hence the highly unorthodox 'back to basics' approach of this paper.

When writing the paper, I had in mind two crucial and interrelated debates that are taking place at present within the United Kingdom actuarial profession: the structure of the examination syllabus; and the future of the profession. Ever since qualifying as a Fellow of the Faculty in 1970, I have watched, with great sadness, the decline of the investment course of reading from a highly interesting and thought-provoking combination of practical textbooks and sessional meetings papers to its present bland dominance by statistical methodologies. Now seems the ideal time to halt this decline and to move back towards good, old-fashioned actuarial wisdom, based on practical experience.

My main message for the future of the profession can be summarised in only five words: "Beware of Greeks carrying gifts". Despite their apparent short-term glitter, I genuinely believe that the current methodologies of financial economics will be of as little long-term benefit to the U.K. actuarial profession as the Trojan horse was to the formerly proud, successful and independent citizens of Troy.

On 25 March 1996, at the Institute of Actuaries, Mr Andrew Smith will present his paper, 'How Actuaries can use Financial Economics'. Although his paper, unlike mine, does not attempt to review established models such as the Capital Asset Pricing Model and the Black-Scholes formula, some observations which he makes from his extensive practical experience support two of the most important themes in my paper: we need to make explicit allowance for the potential irrationality of others; and simplistic linear models can lead to unintelligent, trend-chasing behaviour. In the context of an equilibrium model for the personal lines insurance market, he observes that such a model only admits a market cycle to the extent that it can be explained in terms of rational actions by all individual market participants. It would, therefore, be unable to measure the exploitability of competitor irrationality. Also, in the context of the widely recommended approach of fitting simple time series models to historic asset values, he observes that, in practice, assets which have outperformed in the past are projected to continue to do so in the future, with the result that institutions are encouraged to buy right at the top of the market.

Mr Smith makes two very interesting observations in his concluding remarks. He suggests that the models built by financial economists are now much less unrealistic than previously, in that many of the restrictive assumptions have been relaxed in recent years. In general, I still regard these models as grossly unrealistic, the most blatant deficiencies being: no recognition that financial behaviour is far more complex than implied by the standard assumptions; no recognition of differing levels of behaviour; and the absence of a credible theory of risk.

I have more sympathy for the point of view that he expresses in the final sentence of his paper, namely, that the alternatives to the use of financial economics are to leave some questions unanswered or to build a new solution from scratch. I believe that it is the latter course that we should follow. My paper, while not claiming to be the end, or even the beginning of the end, in this difficult quest, might be regarded by some as the end of the beginning.

My paper is incomplete as a blueprint for a new theoretical framework. I ask you to focus on the essence of my case rather than on the technical merit of its expression.

There is, however, one crucial phrase for which I make no apology. In the very last sentence of the paper I say that, in principle, "the general teachings and methodologies of financial economics are dangerously unsound". Had I only said "unsound", rather than "dangerously unsound", you might have inferred that the conceptual errors within the simplifying assumptions were relatively innocuous in nature. In that case, when considering the case against financial economics, you might have arrived at a not proven verdict, rather than the 'guilty' verdict that I earnestly believe is essential for the future success of the U.K. actuarial profession.

Mr D. O. Forfar, F.F.A. (opening the discussion): Rarely, I think, will we find among the Sessional Papers of the Faculty such a broad landscape set out before us. This is more than a journey through the actuarial countryside, it is a journey through an Academy. We find in the faculties of astronomy, philosophy and psychology, as well as business and finance, not disparate nuggets of knowledge, but common strands in the understanding of human behaviour in the face of risk and uncertainty. The author, in bringing together these common strands, has stimulated our intellects and imagination as he has built his picture of how we, as humans, react to risk and uncertainty.

The matter in question here is a debate between the relative merits of two conceptual frameworks, or paradigms, for the operation of financial markets; on the one hand *financial economics*, and on the other hand what I shall call the *proposed paradigm* of the author. The author leaves us in no doubt as to where he stands; he is a powerful advocate of the *proposed paradigm*, and someone who has very serious reservations about *financial economics*.

The alternatives are broadly as follows. Firstly there is *financial economics*, whose features are:
— markets react immediately to the available information, and only change in response to new information;
— the behaviour of all investors is rational; and
— it is futile to try to outperform the market on a risk-adjusted basis.

Secondly there is the *proposed paradigm*, whose features are:
— there are different types of investor in the market;
— the behaviour of different investors depends on their circumstances, including their wealth;
— the behaviour of different investors depends on their experience and training;
— investors are not prepared to accept risks greater than a certain threshold, and they evaluate risk by the downside measure of an unfavourable result;
— the behaviour of other investors in the market requires to be taken into account; and
— models can be used to identify undervalued or overvalued shares.

I think that it is fair to say that some of the assumptions of *financial economics* have had to be made in order that mathematical development of the subject could take place. Thus, there has been something of a trade-off between assumptions and mathematical tractability, and a compromise has been made. The question is whether the compromise has moved the mathematical model too far away from the real operation of the market.

On the other hand, to build a full mathematical model using the rules of the author's *proposed paradigm* would be very difficult, requiring, it would seem, a feedback loop between the actions of investors and the returns and variability of returns of the individual shares.

As a former investment practitioner, I have to confess something of a bias in favour of the *proposed paradigm*, and would like to make certain observations based on my experience of having worked for some years of my career in the investment department of a life office, and having taken a broad interest in investment matters since then. However, certain of my remarks do recognise certain evidence pointing towards the other side of the debate.

The notion of rational behaviour in investment is an interesting one. What might appear rational to the man in the street can often be far from rational when playing the rules of the investment game. Anyone who has worked in an investment department for only a few days will know that an announcement from a company of increased profits does not mean that the share price will rise. Profits are judged in relation to expectations. If profits fail to satisfy expectations, the share price will almost certainly fall. If the expectation is already in the market, then the news confirming that that expectation has been realised will not change the market. The investor is, therefore, left with the difficult matter of judging what expectations are already discounted, and whether he or she believes

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

reality will exceed or fall short of these expectations. These judgemental matters are not easy, and test the very best of the investment professionals. One professional may conclude that market expectations are over-optimistic, and therefore sell; another may conclude that the market is undervalued, and buy. They may both have satisfied the requirements of rational behaviour, as outlined in ¶24.6, but do different things. It is reasonable to suppose that certain investment managers may make their judgements more skilfully than others, thereby giving them an advantage. If certain investment managers are capable of superior judgement, then this would seem to argue in favour of the *proposed paradigm*.

There can be no-one who has worked in investment who has not chastised himself at some time for a poor investment decision made earlier. At the later date he realises that he had assessed the earlier situation in the wrong way, and that if only he had posed certain questions to himself at the earlier time, then the decision would have been different. He recognises that the ingredients were there, at the earlier date, to enable a better investment decision to have been made, and that he should have been able to see the truth with greater clarity. He learns for the future. This observation would seem to support the *proposed paradigm*.

On the other hand, statistics showing the degree to which investment managers of, for example, unit trusts fail to outperform the market are not encouraging for the *proposed paradigm*. I quote from the *Scotsman* of 24 February 1996: "Virgin Direct commissioned research which examined the performance of actively managed portfolios held in unit trusts over a 20-year period. It found that in only 6 of 121 individual periods analysed did the average performance of actively managed funds better the All-Share Index." Virgin decided, therefore, to track the Index for its Growth PEP. If such analysis raises doubts about the ability of investors through application of equity models and fundamental analysis to identify the subset of shares which are going to outperform the index, then there is evidence pointing in the direction of *financial economics*. The statistics put forward by Jensen seem to be of a similar type. The author, I know, advances arguments which suggest that Jensen's conclusions are not valid. I was left a little uncertain by the author's arguments.

I now turn to the theory of risk. The author is not in favour of variance of the portfolio's investment return as a measure of risk. It would be interesting to know, however, if the framework put forward by Professor Markowitz can, with only minor modification, accommodate within it a much wider definition of risk than the variance of the investment return on the whole portfolio. As I understand it, the essential feature of portfolio selection is that the mathematical model selects the shares and their weightings for the optimal portfolio. *Optimal* can be defined in a number of ways. For tractability of the mathematics, it was initially defined as the minimum variance of the portfolio's investment return, subject to the expected value of the return exceeding a given fixed amount. However, I feel sure that, with the mathematical algorithms now available, the optimal portfolio can be defined as the portfolio with the minimum risk, where the definition of risk is input by the user. In this way it would seem that the definition of risk given by the author could be accommodated. In the formulation given below a possible approach is suggested.

Formulation

Suppose that £1 is invested in n possible shares. The portfolio vector $(x_1, x_2, x_3,..., x_n)$ represents the proportion in each share:

$$\sum_{i=1}^{n} x_i = 1. \tag{1}$$

Vector $(r_1, r_2, r_3,..., r_n)$ represents the random variables which are the investment returns on the shares.

Vector $(\mu_1, \mu_2, \mu_3,..., \mu_n)$ represents the expected investment returns on the shares.

Let r be the random variable which is the portfolio return:

$$r = \sum_{i=1}^{n} x_i r_i.$$

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

We wish to obtain an expected investment return on the portfolio which exceeds a given amount E, say:

$$\sum_{i=1}^{n} x_i \mu_i \geq E. \quad (2)$$

Let $p(r)$ be the density function of r as the $\{x_i\}$ vary.
Let μ be the expected value of:

$$r = \sum_{i=1}^{n} x_i \mu_i .$$

Markowitz 1. Find $\{x_i\}$ to minimise the variance of r, subject to constraints (1) and (2):

$$\text{Variance} = \int_{-\infty}^{\infty} (r-\mu)^2 p(r)dr.$$

Markowitz 2. Find $\{x_i\}$ to minimise the semi-variance of r, subject to constraints (1) and (2):

$$\text{Semi-variance} = \int_{-\infty}^{0} (r-\mu)^2 p(r)dr.$$

Clarkson. Find $\{x_i\}$ to minimise the risk shown below, subject to constraints (1) and (2):

$$\text{Risk} = \int_{-\infty}^{L} (L-r)^a p(r)dr$$

where:
L is the threshold of harm; and
a is the stated parameter.

I now turn to an area where mathematical models of the stock market have proved to be very useful in actuarial work — the forward projection of life offices. The projection of the liabilities of a life office has never posed significant problems, but a proper office projection needs both the assets and the liabilities to be projected simultaneously. This requires simulation of inflation, fixed-interest yields, dividend yields and equity prices into the future. In this way, both assets and liabilities can be projected forward on a mutually consistent basis. The eight interconnected equations devised by Professor Wilkie have proved a very useful tool for undertaking such projections. The primary equation among the set of eight is the equation governing inflation, and this stands independently from the other seven. Should a consensus view emerge that it may be useful to add certain non-linear terms to the inflation equation, along the lines suggested by the author, then I think that I am right in saying that this could be accommodated without disturbing the other equations.

I now turn to the interesting theory of sports put forward by the author. I am sure that there are many here who have personal experience of the sports mentioned, and will have identified with the author's analysis of risk, particularly with the references to ski-ing. I know that I have learned from experience in managing this risk. Usually this experience has taken the form of direct and rather unpleasant contact in a prone position with the hard ice on the ski slope. Perhaps this shows that coping with risk is still very much a human endeavour.

Mr J. Plymen, F.I.A.: As the joint author with Mr Clarkson of a previous paper on the same subject, I wanted to write a parallel paper. The text of my paper has been accepted as a written contribution. I can, therefore, confine my remarks to the vital principles leaving the detailed supporting argument to the written contribution.

In ¶32.4.2 the author says " .. a much better theory, and, moreover, one that is essentually actuarial in nature, is well within our grasp." In my supporting paper I outline the principles for this theory and illustrate how it can be applied in practice to secure the maximum performance with minimum risk.

Every share in the portfolio is continually monitored against the index. Initially the whole portfolio is invested in an index fund, which is used to buy shares only when they are cheap. If there are no cheap shares, the investor stays with the index fund. The investor needs special statistical coverage, something like the Clarkson long-term model, to say whether the shares are cheap or dear, long term. We also suggest that an operational research technique of quality control is used to show whether a share is cheap or dear, short term. The expected return, long term, moves from time to time. The price has a very strong cyclical short-term variation, providing opportunities for profitable trading, given indicators which make sure the timing is right.

In my written contribution I show a chart of Sainsbury's shares over the past five years. I show that we went into Sainsbury's in 1988, when it was relatively cheap. We kept it for about three years and we sold it just before the price turned down. During the period it was held the price over-performed the market by about 16% p.a.

Ideally an investor should always be in over-performing shares. Some shares will over-perform the market for a year. Others, like Sainsbury's, over-perform for three or four years.

The other part of my written contribution is an attack on the Markowitz system — it really is an attack on the Sharpe model. In our previous paper, 'Improving the Performance of Equity Portfolios', (Clarkson & Plymen, 1988) we showed that it suffers from a whole series of basic errors. In the discussion there was no defence of the basic theme that risk equals variance. Even Professor Markowitz now equates risk to semi-variance.

I have found a vital flaw in Markowitz's original set up. It seems rather extraordinary that I have found it and nobody else has seen it over 45 years. Markowitz's measure of risk is wrong. The expected yield allows for all possible risks if it is obtained by investment analysis. If it is obtained by history of prices, the prices in the past have been adjusted for adverse movements, and the expected yield is after risk. There is no power on earth that will enable one to say what the expected yield is before risk. Markowitz was taking the expected yields from price history, but assumed that past price history was not affected by the risk, and in his original method he takes off a further slice of risk, being the variance.

This means that shares are ranked according to (Return−Variance), instead of the expected yield after risk. The undue emphasis given to risk distorts both the share selection process and the calculations based on the efficient frontier. Consequently the method is invalid, both in theory and in practice.

Finally, I make a plea for this situation to be recognised by the Tuition Service and the Board of Examiners.

Mr P. F. Rains, F.I.A.: The paper is of grand design and broad sweep, and sets out to put to rights the fallacies preached by financial economists since their emergence over 40 years ago.

The paper is particularly important to our profession, as it will play a part in deciding whether we follow the work of financial economists or attempt to develop our own approaches and theories independently. In the paper many analogies are drawn with the developments of theories in other scientific fields, particularly astronomy. The author believes that a Copernican revolution is required in financial economics. Rather than providing a better alternative, it is just possible that his approach is like the mother who throws the baby out with the bath water.

The crucial question for actuaries is whether the author's approach will lead to a new age of enlightenment or a new dark age for the actuarial profession in investment, accelerating our declining influence in investment matters.

I first consider the dynamic equilibrium model. The fact that securities fluctuate around the mean absolute deviation is almost a truism. It is independent of whether markets are efficient or inefficient. Then, to show that his model works, the author needs to show that superior investors exist, and can make money using his approach. He has not shown that in his paper — or he has certainly not proved it in a statistical sense.

One obtains value from the author's approach only if one can identify and use the valuation model. A widely published and accepted valuation model seems a sure route to ensuring that it has a limited lifespan in terms of creating outperformance. That is not to say that financial economists believe that inefficiencies do not exist. Many papers from financial economists have shown that they do. Markowitz, himself, has developed a model based on the existence of market inefficiencies.

I now turn to the author's predictions for his model and theory. Starting with his first prediction, he has yet to show that the model users have high returns, allowing for the possibility of luck and for data mining.

Turning to the second prediction, this can be rephrased as: if one can identify undervalued shares, one can reduce risk. However, this can be predicted within a mean variance framework, and has been a subject of many papers by financial economists using the downside risk measure.

The third prediction, despite what the author says, is consistent with modern portfolio theory. This says that one should expect, *ex-ante*, that investments with high risks have higher expected returns, *ex-post* you should expect the more volatile investments to have a higher dispersion of returns. That is just what his table shows. The fourth prediction of greater share volatility with greater earnings uncertainty is uncontroversial, and cannot be put forward as a major new implication of this theory.

I now turn to investment risk. Downside risk is a summary statistic that might be appropriate for some investors. Other investors might need criteria which give a fuller description of the return distribution relative to their return requirements. It seems unlikely that a single summary statistic of risk will provide sufficient detail in all cases.

Also, I should like to attempt to give a counter example to the author's risk criteria. Let us suppose that we have an investor with £110 who requires at least £100 in a year's time. Let us also suppose that he can invest in cash or in a risky investment, which we will call equities, and that this risky investment has a higher return than cash. Let us also suppose that he can purchase derivatives on the risky asset. What is the optimal asset distribution for the investor using the author's criteria? It turns out to be an investment in cash which has a value of £100 in a year's time, with the rest invested in very out-of-the-money call options on equities. This has no downside risk and provides the highest exposure to equities, as they rise, and therefore the highest expected return. However, the very high expected returns from this portfolio are rarely achieved, and the portfolio would be chosen by few, if any, investors.

Thus, the essential elements of the author's new approach, where they differ from recent research of financial economists, are either uncontroversial or unproven, and possibly also flawed. The paper has many interesting insights, but these can be looked on as adjuncts to, or minor variations of, the more recent work of financial economists rather than a refutation of the whole financial economics approach.

It may be true that the author has played a part in reforming the practice of financial economists in recent years, by encouraging the more prevalent use of the downside risk measure. I hope that he will continue to look for further refinements of a similar kind.

Mr A. D. Smith: I am speaking as someone who uses financial economics on a daily basis to advise insurance and banking clients. I am only too aware of the shortcomings in my techniques, but they do seem to be the best that are currently available. I would be prepared to sacrifice a few sacred cows, such as market efficiency or utility theory, if I could find something better which would give deeper insight and lead to more profitable decisions.

Having read the paper, I have in mind a number of areas where current practice can be improved. I wish to comment on three elements of the paper, these being: downside risk; dynamic equilibrium; and the impact of irrational investors.

I am sympathetic to the author's concept of downside risk. It seems more logical to me than standard deviation as a risk measure. For example, spending £1 on the National Lottery entails roughly the same standard deviation as betting £1,000 on a coin toss. To my mind the coin toss has a much higher risk.

I hoped to see reference to the 1977 paper, 'Capital Market Equilibrium in a Mean-Lower Partial Moment Framework', by V. S. Bawa & B. Lindenberg, published in the *Journal of Financial Economics*. This paper suggests precisely the downside risk definition which the author has adopted in ¶10.7.2, and explores how the world would behave if such risk measures were adopted by all investors. The author has also explored this issue and, in ¶17.5.1, he claims that the downside risk

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

concept explains many types of behaviour that are inconsistent with expected utility theory. This contrasts with Fishburn's work in 1977 ('Mean-Risk Analysis with Risk Associated with Below Market Returns', by P. Fishburn, published in the *American Economic Review*), who demonstrates how a more general downside risk concept may be accommodated within an explicit utility framework. In the light of Fishburn's work, the stark contrast between utility theory and downside risk, outlined in Section 11.5, deserves some explanation.

The author cites two normal distributions, claiming that one lies to the right of the other. I believe that this claim is mistaken. However, he illustrates his point using two approximating binomial distributions which do satisfy his claim. He creates a clever illusion by testing his own method on the binomial distributions, while testing expected utility on the normal distributions. Unsurprisingly, the results are inconsistent. The problem lies with the approximations he had used. If one applies utility theory to the approximating distributions, the conclusion would reinforce those obtained by the author's own method. On the other hand, as the cumulative distributions do intersect, it follows that decisions based on semi-deviation, with a suitably small value of L, would favour the distribution with lower variance, which is consistent with the textbook.

So, what can we conclude from the author's downside risk concept? Certainly we cannot reject utility theory, which is a generalisation, without rejecting the author's special case outlined in the paper. However, it is evident that mean-variance analysis has some significant problems which have been addressed at length in the financial economics literature.

Turning to dynamic equilibrium; classical financial economics considers the outcome if investors correctly implement rationally formulated objectives. The author has convincingly argued that investors do not act in this fashion, and, consequently, the conclusions of equilibrium theory, particularly the relationship between risk and return, must be open to question.

The author has laid out his own behavioural theory in great detail. It must surely be of interest to consider how markets would behave if investors conform to the author's description rather than the rational behaviour of the classical theory. An economist's approach would be to build an equilibrium model where investors followed their possibly unintelligent objectives and prices equated supply and demand.

The author's own equilibrium model is given in Section 9.8. However, I am not sure that the suggested approach is consistent with equilibrium in the usual economic sense. I do not think that the application of Newtonian dynamics to asset prices is consistent with supply and demand considerations. For example, when prices have peaked and are starting to fall, type-B investors are trend chasing and thus selling, while types-O, A and R are selling because the price is dear. If all four investor classes are selling shares, who are they selling them to?

The author makes a number of claims for his dynamic equilibrium wave model that merits further investigation. For example, I noticed that the dynamic equilibrium model arises from a second order linear differential equation. How then can it be consistent with non-linear chaos theory, as claimed in Section 23.1?

In ¶27.3.1 the author endorses Mandelbrot's theory, which produces jumps and fat tails. Where are the fat tails and jumps in a sine wave?

By the same token, in ¶27.3.2, it is claimed that the dynamic equilibrium model is consistent with Mandelbrot's scaling principle; but surely the distribution of an increment over a time lag depends on the size of the lag relative to the wave length, in violation of Mandelbrot's condition.

I do not agree that the scaling principle is totally inconsistent with the standard assumptions of financial economics. The Guassian random walk is one example of a process which satisfies the scaling principle on the impact of irrational investors.

The author has considered many assumptions and models produced by financial economists. In many cases he has reasonable doubts regarding the validity of assumptions. This begs the question: does it matter? In other words, are the models of financial economics robust departures from the assumptions?

My particular concerns are in the areas of managing insurance companies and other financial institutions. My usual method takes a broad brush approach, assuming that investment returns follow a suitable index. Can I continue to get away with this macro view or should I drill down more deeply into the asset management process?

The conclusion that I come to is that, while market return distributions will continue to be the major ingredients on the asset side for asset-liability studies, manager skill and execution ought also to be important. That should not conflict with wider considerations of market efficiency when there are transaction costs.

I confess that the asset liability methodologies that I use do not allow for an assessment of skill, and, therefore, could miss important aspects on the asset side. I do not accept that the current practice is dangerously unsound, but there is certainly scope to improve the economic theory to take account of the features which the author has pointed out.

The author's claims to have disproved financial economics are mistaken. What he has done is to demolish a straw man of his own creation, a caricature of financial economics which I do not recognise. However, if we strip out what I see as an unwarranted assault on work which is, in the main, sound, there remains a fascinating account of market behaviour, gleaned over a quarter of a century's practical experience, which we should not dismiss lightly. While I do not agree with every point expressed in the paper, I consider myself richer for having been exposed to it.

Dr A. J. G. Cairns, F.F.A.: Before I discuss some sections in detail, I should like to make a general comment. There are several instances in the paper where a single dataset is regarded as proof that the author is right and the financial economists are wrong. In other places extreme examples are used. I have some sympathy with this approach, but only when it is used to help the understanding of an idea. I absolutely disagree that such examples can be used to prove a particular point.

One example is in Section 11.5, in which the author attempts to prove the value of downside risk. Now, I agree that the downside measure of risk is reasonable, and the example given in ¶11.5.1 is apparently quite convincing. However, I tried to match the result by considering what I regard to be plausible combinations of expected return and volatility over, say, a one-year period. For example, we could have a risk-free asset offering a return of 6%, a medium risk asset giving a mean return of 8% and a standard deviation of 10%, and a higher risk asset with a mean of 10% and a standard deviation of 20%. In this case the outcome was quite the reverse of that in ¶11.5.1. The higher variance asset was about twice as risky as the lower variance asset, using the author's measure of risk. This was true for all reasonable values of L (say, from 0 up to 20%, and did not depend in any way upon whether or not the exact formula or a Binomial approximation had been used). No matter how hard I tried, I could not come close to replicating the author's results (using sensible means and variances). On the contrary, with any reasonable pairing, higher variance assets always turned out to have a higher risk, which is what we would expect.

So, why was I unable to replicate the author's result? A simple interpretation of the author's example is to take the numbers as percentage returns. Both assets then have relatively low variances compared to equities; but, remarkably, the second asset offers a whole 1% extra return for only a tiny increase in variance. So, it is hardly surprising that this asset seems to be much more attractive. There is no disputing the fact that the values given in ¶11.5.1 are correct in the numerical sense; it is the extreme choice of input parameter values which I object to, particularly since it leads to an incorrect conclusion.

The other section that I would like to discuss is Section 12, in which the author describes his inflation model. The model is perhaps not unreasonable. However, I found his arguments in favour of his model and against, for example, the Wilkie model so irrational that it is actually very difficult to argue why his approach to this problem is wrong. Part of the problem is that here the author, as in other places, has failed to use such things as mathematics and specific parameter values. For example, in ¶12.9.2 the author states: "For the other parameters, I used 'actuarial judgement' rather than any conventional statistical approach". To me actuarial judgement should represent two things. First, parameters should be chosen with reference to past experience. Second, a model should, in this context, make economic sense. I have a feeling that when actuarial judgement goes further, too much emphasis is likely to be placed on recent experience. If the author feels that parameter values need to be modified to reflect recent structural changes, then surely he should admit that his model is inadequate, since it does not admit the possibility of similar structural changes in the future. Presumably the model would have to add in further random shocks which alter parameter values, and, quite frankly, this would make the model even more unpalatable.

In ¶12.10.3 the author quotes from Geoghegan *et al.* (1992), pointing out that their report supports his conclusion that his model provides a superior fit. This is hardly surprising, since you are bound to get a better fit if you include extra parameters.

In Section 12.14 the author notes that his model cannot be assessed on the basis of statistical tests. This is nonsense. If the model has been properly formulated, then it must be possible to fit the model statistically to past data. It may not be easy to carry out this task, but it is possible.

Mr S. J. Green, F.I.A.: In his latest book, *Money, Credit and Asset Prices* (1994), Professor Pepper says that monetary systems are rarely in equilibrium. Of course classical economic theory is all about equilibrium. It is all based upon financial markets or economies either being in equilibrium or tending to equilibrium. Frankly, I do not think that they are.

It is part of classical economic theory that one has a system which is in equilibrium, and then one has an external shock which knocks it out of equilibrium. It fights very hard to get back into equilibrium. This is the theory upon which classical economics is based. It totally ignores modern theories of complexity that say that the systems themselves have within themselves the power to generate those shocks. The shocks are not external.

In the late 1940s and early 1950s there was a children's toy called a Tippy top, which was a spinning top. You spun it on its base and then all of a sudden, but for no apparent reason it would turn and spin on its point. The Tippy top is very much like the financial markets; they are apparently in equilibrium, but they are not. They have within themselves, at any time, the power to shock.

I have an analogy, and thought that a financial economist is like a physicist looking at four glasses of water. One of them is Highland Spring; another is distilled water; a third one might come out of the sea at Leith; the fourth one comes from the Dead Sea. To the physicist they are all just water, because they all have the same properties: they all boil; they all freeze; ultimately they all evaporate; they all pour in exactly the same way. However, if you go to a fireman, or a fisherman, or a diner, or a motor mechanic, then each glass has water with different properties. The practitioner's view is totally different to the view of the financial economist. We have a lot to learn from financial economists, but we should not accept what they suggest hook, line and sinker, because they ignore what is going on in the market. It is very difficult for practitioners to apply financial economics without a considerable amount of subjective judgement.

The opener brought up the illustration of unit trusts which failed to out-perform the index, and used this as an example of how the author's theories were wrong and financial economics were right. Together with Day, Plymen and I worked for a couple of years on a paper which we produced at one of the AFIR conferences. In order to do so, we examined all the available quant models. There was not a single quant model which we could find which out-performed the index. Those which came quite close contained a considerable amount of subjective judgement. So, financial economics and modern portfolio theory, by themselves, are not of significant use. It is certainly wrong that our students are being brought up to apply them willy nilly. If they are applied with true actuarial judgement, then they have something to offer.

Stochastic models are fine if they are used properly, but they are not there for forecasting what is going to happen. They will forecast what could happen in certain circumstances. In order to forecast what is going to happen, you need subjective input. You need to go back and do what we were taught to do: examine the past, look at the present and predict the future from them.

The paper has reopened the debate, and it has done it in a way which, for those whose minds are not totally closed, will provoke more thought as to how we should apply financial economics.

Mr J. E. Paterson, F.F.A.: I agree with the author that, from about the time I came into the investment field, the antics of financial economists have become increasingly bizarre. This does not mean that financial economics is incapable of making a contribution; and it does not mean that all practitioners of financial economics are snake-oil salesmen.

Much of the trouble with financial economics comes from putting too great an emphasis on statistics, which, in reality, is of little use. It is not unreasonable to say that what happens in the world is determined by realisations of probability distributions. Probably it is. The trouble in the investment field is that you only ever get one realisation of any distribution. Every year's investment return in, say, the U.K. equity market may be a realisation of a probability distribution, but if next year's is a different probability distribution from this year's, and last year's was a different one again, how do you expect to be able to get much of value out of statistics? In this area it is important to realise that numbers are like alcohol — a good servant, but a poor master.

Turning to a comment of an earlier speaker, I agree that a frequent result of the downside risk approach would be that, once you have safeguarded the minimum required return, then there is a very strong drive towards disregarding any risk-to-return trade-off in other outcomes, and just going for the highest return. That cannot be the perfect answer, yet the author's emphasis on downside risk is a step forward. I think that it was Isaac Newton who said, "I have seen so far because I stood on the shoulders of giants." Science must usually proceed on the basis of incremental improvements on what went before. Perhaps once in a few hundred years you get a Newton or an Einstein who can make a massive leap. Most of the time we must be grateful for the people who can help us take small steps in the right direction.

The author has made a start, in a way that he admits is largely destructive. I am afraid that the foundations of what is habitually done in the attempted application of financial economics to investment are pretty spurious. The driving forces have been that the models used are fairly tractable in mathematical terms, and that the computations required can actually be done with the aid of modern computing power. However, with the theory, as with the calculations, the old adage holds: garbage in; garbage out. Not nearly enough thought is being put into the fundamentals of what the numbers mean, and to the fact that the vagaries of human behaviour and fashion play an enormous part in determining what people actually do.

This 'big picture' problem is something that should be thrashed out thoroughly. We have heard a financial economist trying to rubbish the paper with as much enthusiasm as the author has been trying to rubbish much of financial economics.

This is probably the most important investment topic that has been discussed in the Faculty in all the time that I have been attending meetings. Let us keep the gloves off and fight dirty on all sides.

Dr D. C. Bowie, F.F.A.: While I appreciate the immense effort that the author has expended in providing us with a most enjoyable paper and an alternative approach to investment risk, I regret that his antipathy towards financial theory has meant that he has not investigated it beyond a rather basic level. In particular, much of his venom is directed at the earliest and least sophisticated of the financial models, and these are, of course, the most unrealistic. This is the nature of model building.

It is my view that financial economics does provide great insight into how financial markets might work; that financial economics does provide a very elegant and plausible set of links between portfolio theory, asset pricing and corporate policy; and I think that financial economics must be one of the best endowed areas of research in terms of empirical testing, model adaptations and modifications. The inconsistencies brought to light in empirical testing have not been ignored. The unrealistic assumptions required for the basic versions of the models are not actually thought to be gospel; and research in financial economics did not grind to a halt after 1959.

The author dislikes variance as a measure of risk, and provides us with an alternative downside measure in ¶10.1.1, which we can recognise as being a lower partial moment of the distribution of security returns. That measure appears very sensible, and may well turn out to be a widely adopted measure. Financial economists would probably agree with the author, having developed both portfolio selection and asset pricing theories based on downside risk, starting with work by Bawa & Lindenberg in 1977. Subsequent work has shown that utility theory is consistent with this downside measure of risk and that there is not much practical or theoretical difference between the two models.

Other researchers in financial economics, for example Kraus & Litzenberger (1976), have derived portfolio selection theories and asset pricing models in mean-variance-skewness space, where positive skewness is counted as a desirable feature of the asset's return distribution. Indeed, it is theoretically feasible to use all the moments of security return distributions for selecting portfolios and pricing assets, although the practical advantage of doing so is debatable.

The early theories in financial economics did, indeed, have the most horrendously unrealistic assumptions about equal tax rates, perfect liquidity, instantaneously available information, uniformly intelligent or unintelligent investors, etc. Most financial economists would agree entirely with the author that any model based on those premises is unlikely to produce anything even vaguely sensible. Financial economists have not ignored that incompatibility; a multitude of researchers in the late 1960s and early 1970s (people such as Black (1972), Mayers (1973), Brennan (1970, 1971), Merton (1973), Lintner (1969), Dyl (1975), etc.) extended the basic models to allow for differential tax rates, the absence of a risk free asset, the existence of unmarketable assets, uncertain inflation, short sales and many other realistic factors ignored in the original formulation.

Financial economists would not be uncomfortable with the author's categories of investor based on intelligence and/or information — Lintner (1969) and Grossman & Stiglitz (1976) (among many others) have created consistent portfolio selection models and equilibrium models for asset pricing, based on inhomogeneous information and expectations.

The major point that struck me in reading the paper is that financial economists would agree about many of the author's observations regarding unrealistic assumptions in the models. Many of them would agree sufficiently to go off to try to create alternatives which were not subject to those criticisms. However, most of the author's criticisms have already been dealt with in the mainstream financial economics literature. There does not exist an identity between a bumbling mean-variance addict and a mainstream financial economist.

In Secton 9.5 the author proposes a Turing test for market efficiency. His test is actually a test for predictability, not a test for efficiency. The concept of efficiency, as proposed by financial economists since the late 1960s, is based on the speed at which information is priced into securities. It does not necessarily have anything to do with predictability, and, in fact, it is theoretically possible to have a perfectly predictable market which is perfectly efficient. It is absolutely and categorically untrue that mainstream financial economists (after about 1970) believe that stock prices follow a random walk — we must be careful to distinguish real financial economics from the early interpretations of early theories.

The author's entertaining and controversial paper does, by its nature, elicit strong responses. I do not agree with his perception of financial economics, but I did welcome the opportunity to review the very important basic considerations which underpin my acceptance of much of financial economics. I think that the author would have engendered a more positive response to his ideas if he had contrasted his work more carefully and less aggressively with modern financial economics and academic models. In particular, I think that the periodic table approach to theory-worthiness is artificial and counter-productive. I would further recommend a more comprehensive and less biased review of what does exist in financial economics before dismissing it.

REFERENCES

BAWA, V. & LINDENBERG, E. (1977). Capital market equilibrium in a mean-lower partial moment framework. *Journal of Financial Economics*, **5**, No 2, 189–200.

BLACK, F. (1972). Capital market equilibrium with restricted borrowing. *Journal of Business*, **45**, No 3, 444–455.

BRENNAN, M.J. (1970). Taxes, market valuation and corporate financial policy. *National Tax Journal*, **23**, No 4, 417–427.

BRENNAN, M.J. (1971). Capital market equilibrium with divergent borrowing and lending rates. *Journal of Financial and Quantative Analysis*, **6**, No 5, 1197–1205.

DYL, E. (1975). Negative betas: the attractions of selling short. *Journal of Portfolio Management*, **1**, No 3, 1174–1176.

GROSSMAN, S. & STIGLITZ, J. (1976). Information and competitive price systems. *American Economic Review*, **66**, No 2, 246–253.

KRAUS, A. & LINTZENBERGER, R. (1976). Skewness preference and the valuation of risk assets. *Journal of Finance*, **31**, No 4, 1085–1100.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

LINTNER, J. (1969). The aggregation of investors' diverse judgments and preferences in purely competitive security markets. *Journal of Financial and Quantative Analysis*, **4**, No 4, 347–400.
MAYERS, D. (1973). Non marketable assets and the determination of capital asset prices in the absence of a riskless asset. *Journal of Business*, **46**, No 2, 258–267.
MERTON, R. (1973). An intertemporal capital asset pricing model. *Econometrica*, **41**, No 5, 867–887.

Dr A. S. Macdonald, F.F.A.: The author is, perhaps, the leading proponent of the view that the actuary should not add financial economics to his or her toolkit and that actuarial science should seek its own way forward. This paper is the strongest expression of this view so far. The subject matter of financial economics is, by its nature, mathematical and statistical. How can economic phenomena be modelled, and how well do such models stand up against real data? In ¶1.16.1 the author, more or less, abandons mathematics in his paper. It soon becomes clear that he has no desire to allow the precision of mathematical argument to stand in the way of his assaults on his chosen targets. Given his literary preference, those mathematical arguments that he does deploy are of key importance in judging the worth of his views as a whole. On this evidence, unfortunately, we are justified in treating the author's thesis with rather more than caution.

When the author attempts to cast doubt on stochastic models he does so from a standpoint which would barely be recognised by anyone with some knowledge in this area. This is important. The author describes his targets for the benefit of an actuarial audience, not all of whom will have extensive knowledge of these fields; but his descriptions are his and his alone, and they are limited by his own historical perspective. To anyone who does not bring to their reading of this paper their own knowledge of its subject matter, I would say; "Look elsewhere to find out what financial economics is about. You will not find it here".

It is not only in his descriptions that the author is at fault; much of the content is technically deficient. An important example is his treatment of Crusoe's dilemma — important because he bases much upon it. In Section 2.13 he puts forward a stochastic model for Crusoe's chances of solving his puzzle in time. It turns out to be a compound distribution. The time taken to solve a problem has a Normal distribution whose mean value is itself random. This models the situation where it is not possible to tell, at first glance, just how difficult the problem will be, and it is, therefore, appropriate for use if that particular piece of information is not available a priori.

In ¶2.14.1 Crusoe is confused because his common sense tells him that he is likely to solve the problem in 15 minutes. There are two possibilities. First, suppose Crusoe is correct and he is able to tell, a priori, that this is a puzzle that he can probably solve within 15 minutes. Then the given probability model is the wrong choice, because it is a model of the decision process without that piece of a priori knowledge. It is not the use of a stochastic model which is at fault, it is the choice of a model which does not reflect the true situation. The second possibility is that Crusoe is wrong, and that his a priori assessment of the difficulty of the problem is unreliable. In that case, even an approximate model might be a valuable aid to his judgement.

From logical error the author proceeds to more serious crime. He has chosen an inappropriate stochastic model and overlooked other models which might be more appropriate. Then he points to an absurd outcome and tells us, in ¶2.15, that this is an example of "generally accepted applications of stochastic models". Generally accepted by whom? As an example of an unacceptable misapplication and misuse of a stochastic model, his example will serve; as an example of how someone well versed in stochastic models would proceed, it is hopelessly inaccurate.

We gain some insight into some misconceptions about the modelling process in Section 12.4, in which the author expresses again his preference for plain English and says, "My Robinson Crusoe example shows how this problem can be mitigated to some extent by translating the initial statistical formulation into something more recognisable to numerate professionals." Numerate professionals are not the problem, it is the others that we ought to worry about. By all means describe the ideas behind models in plain English, but to remove mathematics from the discussion of quantitative or quantifiable matters has one outcome only: to shield speculation from scrutiny.

Given how little mathematics there is in the paper, it is remarkable how much of it suffers from inaccuracy of one sort or another. Each example might be trivial, taken in isolation, but in a paper which would have us reconsider the foundations of a technical subject, the reader is entitled to expect a high order of competence in the details, or is entitled to view the major arguments with suspicion if they are supported inadequately.

Here are some examples. In ¶25.2.5 the author says: "If we make the usual statistical assumptions relating to independence and the relevance of the Central Limit Theorem (generally known as the 'Law of Large Numbers')". The Central Limit Theorem and the Law of Large Numbers are *not* the same. In ¶20.2.3 the author makes the statement that A,B,C,D,E and F together imply Z, and says, "From my training as a pure mathematician, I observe that any inference of 'strong level' efficiency" — that is, Z — "is invalid unless A, B, C, D, E and F are all true." My training leads me to state that the given statement says no such thing. It is the opposite: that Z implies A, B, D, E, F, that would have that result. Also in ¶20.2.3 the author says, "the probability of A, B, C, D, E, and F, all being true is the product of the probabilities of each of these independent statements being true." Dressing up a nonsensical statement in the language of probability does not make it sensible, and this is just gibberish. Then, in ¶27.3.2, the author states that Mandelbrot's stable distributions are, "totally inconsistent with the standard assumptions of financial economists". They are not. The Normal distribution is simply a special case.

The author's main claim to innovation is the dynamic equilibrium model of Section 9.8. From this point on the paper becomes increasingly bizarre. Having dismissed (in Section 9.7) a suggestion by Cootner from 1962 (and was this the most recent contribution by a financial economist that the author could find?), a model is introduced which turns out to be simple harmonic motion plus a, presumably, random error term. The simplicity of the model results from some assumptions every bit as heroic as efficient markets and with less chance of being tested. The main feature of the model assumptions is that they lead to a very simple, mathematically tractable, result; yet such tractability is one of the main criticisms of the work of financial economists.

In Section 9.9 some predictions are discussed. We are not told how these follow from the model. We are relieved to learn from the first two that Clarkson in 1996 agrees with Clarkson & Plymen in 1988. The third is illustrated with one small data set and contrasted with modern portfolio theory (MPT).

One result as predicted does not prove a theory, although one counter to prediction may disprove it. The author states that the result is inconsistent with the teaching of financial economics; namely that low expected returns are associated with low variance. His data, however, relate to actual returns, not expected returns. Drawing conclusions about the latter from a sample of the former requires statistical care. I would also point out that the association of low returns with low variance holds for portfolios on the efficient frontier, and need not hold in general. Unless the author has verified that his 100 portfolios were on an efficient frontier, there is no evidence against MPT whatsoever. Unfortunately, this is not an isolated example, and, as has been mentioned before, his fourth prediction is also no test of veracity.

The specification and description of the model are seriously inadequate. For example, if there are only two classes of investor, then, when one is buying the other must be selling. If, then, intelligent investors make more money, they must squeeze out the less intelligent investors. They will certainly not keep returning to the same position *ad infinitum*. This does not appear to be a model of a market in which investors actually buy and sell shares. Moreover, the real market does not display simple harmonic motion, so an analysis of the error term is of crucial importance. This term is not even specified. Let us be absolutely clear, the author has not specified, described or tested his model even remotely adequately.

It is impossible to comment on all the points raised by this paper, so I risk leaving the impression that I have criticised a small part of the paper, but that I find less fault in the larger part. In order to dispel any such impression, let me say that I mean my comments to apply to the whole paper. The examples are not atypical of the paper as a whole. It shows throughout a failure to comprehend the principles and judicious use of mathematical models.

In a discussion at the Faculty in 1977 one speaker concluded, "I think that it is a pity that a paper such as this should be presented at such a respected forum as the Faculty on the basis of such shallow research. Views expressed in this hall tend to be reported and are sometimes taken as fairly authoritative statements of the view of the profession. It would be unfortunate if tonight's paper is taken as such." I echo that speaker's sentiments.

Professor A. D. Wilkie, F.F.A., F.I.A.: The author says, in ¶20.3.1: "I regard Peters (1991) as one of the most important books on capital market theory since Markowitz", and then he explains why. I have read the book by Peters, *Chaos and Order in the Capital Markets*, and have recently obtained a second book by him, *Fractal Market Analysis*.

Peters presents quite a lot of statistics using an interesting concept 'range versus scale analysis' (RS analysis). However, he asserts that statisticians only know about normal distributions; only deal with uncorrelated variables; only deal with linear models; and, since the RS analysis shows that certain share prices have not moved in pure random walks, they, therefore, are generated by a non-linear chaotic model, using fractional Brownian motion and using Lévy-Pareto stable distributions. I doubt whether he understands what he is talking about. It is quite clear that he does not know the difference between the density function and the characteristic function of any distribution, even the normal distribution. In his second book it is also quite clear that he does not understand either Gaussian distributions or normal distributions, because he thinks that they are different, whereas they are just different words for the same thing.

However, they are both entertaining, readable books, and they have some bright ideas in them. But then *The Hitchhiker's Guide to the Galaxy* is an entertaining, readable book, but, as a text book in astronomy, it stands up at about the same standard as this paper on financial economics.

I have been depressed by the discussion. It is, I am afraid, a dialogue of the deaf, between those who know something about statistics, about financial economics and about mathematics, and those who do not. I am really very discouraged that the Faculty should have brought forward such a paper.

Professor R. S. Clarkson, F.F.A. (replying): I had anticipated a lively discussion with a considerable degree of polarisation towards one side or the other, and this expectation has proved reasonably accurate. I now respond to some of the more important issues raised by the opener and by Mr Smith, both of whom, of necessity, have been rather more exposed than most in recent weeks and months to the subject matter of this paper.

The opener suggested that it would be very difficult to develop my approach into a full mathematical model. I agree that a feedback loop is required between the actions of investors and the returns on individual shares; this is precisely the 'predictor-corrector' approach that I develop in ¶24.2.3. However, human judgement rather than formal mathematics is probably the best way forward, and a good example of the very different practical judgements that have to be made is the reaction of general insurance shares to an increased level of claims. After a 'normal' hurricane, share prices might fall. However, an exceptionally severe natural disaster, causing very serious losses, could trigger the general expectation that the underwriting cycle will change, with all companies — previously constrained by competitive pressures — now able to put their premium rates up. What might have been thought of as bad news is interpreted, through complex interactions of human behaviour, as good news.

The opener raised a very interesting point about whether or not we can extend and improve the Markowitz approach by using input from the user. The problem I see here relates to what I regard as one of the weakest features of financial economics, namely the implicit assumption that all required probabilities are known with absolute certainty. The Markowitz value of risk is correct only if all the assumptions regarding the future are correct, which will seldom be the case, and indeed, as long ago as November 1953, in the discussion on two investment papers, the then President of the Faculty, R.L.Gwilt, drew attention to the dangers of assessing risk simply in terms of our central assumptions as to the future:

> "If you will forgive a brief presidential platitude, I think myself it is wise in the investment of the assets of a life office to follow a policy of moderation and restraint, and to resist any temptation to strive for a spectacular profit if there is any possibility of a serious loss, if one's assumptions should prove to be wrong."

The opener had more sympathy than other speakers, particularly Dr Cairns, for the possibility of generalising Professor Wilkie's stochastic model along the lines I suggest to make it a truly non-linear model. The main problem that I see with any model based on an ARIMA approach is the speed of convergence to expected or central values. Variations about the central yield values that I obtain for U.K. equities and fixed-interest securities seem to indicate convergence to the central values within two, or, at most three, years, after which the behaviour is fairly random. A possible clue here is a change in the value of the Hurst exponent, despite the criticism made of this diagnostic statistic. As well as introducing non-linearity into the inflation series, I believe that we need to look very carefully at whether the rates of convergence of the other equations within the Wilkie model are correct. I suspect that they might be correct, on average, over a very long period, but this could lead to highly unsatisfactory results over shorter projection periods of up to, say, five years.

Both Mr Smith and Dr Bowie were critical for my not mentioning the work of Bawa & Lindenberg, who, they say, had derived precisely the same framework for downside risk as long ago as 1977. My recollection is that the measure of risk that Bawa & Lindenberg use is actually a lower partial moment. This is far less general than the measure of risk that I describe under Axiom 2 in ¶10.9.1, and, indeed, it corresponds to the special case function that I describe under Axiom 7, but with the risk aversion parameter unrealistically restricted to either 1 or 2 rather than being allowed to take any value greater than 1. More importantly, Bawa & Lindenberg do not incorporate the explicit upper level of acceptable risk required for what I call rational behaviour, and hence their downside risk framework is quite different from mine. Furthermore, they build up their model within the context of utility theory, which I criticise severely.

It is, perhaps, worth mentioning that, despite there having been many hostile comments on other aspects of my paper, no-one has disagreed with my criticisms of utility theory. If you go through the papers in the *Journal of Financial Economics,* and discard any that make use of expected utility, there are not many left. I stand by the general classification of utility theory that I adopt in my periodic table. Because of the linearity assumptions, there are levels of sensible behaviour with which utility theory cannot cope, and that is why I suggest that we not only need a downside approach to risk, but we also need a specific maximum level of risk.

I cannot let Mr Smith, eloquently supported by Dr Cairns, get away with their attempted demolition of my flagship numerical example on risk in Section 11.5. If they are unhappy with my general criticisms of current approaches to risk, it is quite natural that they will look for a flagship example and try to sink it; but I fear that they have allowed narrow statistical considerations to knock them off the path of what I would call actuarial common sense.

I say, earlier in the paper, that the main emphasis within actuarial mathematics should be on finding, to a known degree of accuracy and using available numerical data, practical solutions to real world business problems. That is precisely why I introduced the binomial approximation as an eminently sensible practical expedient. It allows you to calculate my measure of downside risk on the two normal distributions very easily, without having to look up any statistical tables and without requiring any detailed knowledge of statistics. There is, however, an even quicker way of calculating these downside risk values for the two normal distributions, and that is to use the risk commutation function approach that I describe in ¶10.8. I used this approach to check the numbers that appear in the paper. As regards numerical data, all that you need is the single page of risk commutation functions for normal distributions that appears in the Appendix to my 1989 paper, 'The Measurement of Investment Risk'. Using these commutation functions, in about two minutes you can obtain values for the final ratio of 2.05, 2.18, 2.29, 2.39. Since these all agree to two significant figures with the values that appear in the paper, I, for one, regard the technical objections to my flagship risk example as overruled.

Mr Smith raised a most interesting question: if you have different classes of investor and they are all trying to sell at the same time, who is buying? The answer is staring us in the face; no-one! When this disequilibrium becomes extreme, prices go into freefall until they come back to what I could call a central value based on long-term fundamentals. Financial economists euphemistically describe this as excess volatility, but those who were involved in the real financial world around the time of the October 1987 'Crash' will realise how terrifying such behaviour can be. This type of behaviour was described by Mr George Soros as a dangerous example of systemic instability in the financial markets, but it does not exist according to financial economists.

Mr Smith, Dr Bowie and Dr Macdonald all criticised my Dynamic Equilibrium Model. I agree that further development is required, but I suggest that we have to walk before we can run, and I am merely putting this model forward as a less unrealistic theoretical framework than the Capital Asset Pricing Model to make financial economists stop and think about some of their assumptions.

A general theme which ran through many of the contributions was that financial economists have worked hard in recent years to remove the simplifying assumptions that could, perhaps, have been criticised for leading to unrealistic models 20 to 30 years ago. As I said earlier, the crucial consideration is whether or not we can accept utility theory as the mathematical cornerstone. Having researched the philosophical underpinnings of utility theory in some depth, and having attended an international conference on the subject two years ago, I stand by the criticisms made in the paper. In short, I do not accept that the simplifying assumptions in relation either to utility theory or to other aspects of real life behaviour have been reviewed in sufficient depth, and I regard the more and more complex models that have been introduced in recent years as still being inconsistent with observed behaviour in the real financial world.

The President (Mr G. M. Murray, C.B.E., F.F.A.): There are clearly two bodies of opinion, and, to a certain extent, there has been wasted effort in attacking the other side. I feel that we should all acknowledge that we are all dealing with genuine uncertainty and future uncertainty. We are all trying to solve the same problems, even though we are coming at them from different angles, or, in some cases, from opposite ends of the spectrum.

I hope that everyone will join me in thanking the author for his efforts in producing the paper, and in providing an enjoyable, interesting and exciting debate.

WRITTEN CONTRIBUTIONS

Dr M. R. Hardy, F.I.A.: I would have made this contribution orally at the meeting, had I been able to attend. It should, therefore, be read as such. The author's opinions of financial economics, as expressed in this paper, contain a fundamental flaw in his reasoning. He rejects many of the results and conclusions of financial economics because he does not believe that all the assumptions which led to those results hold in the real world. This is a misunderstanding of the modelling process and a misunderstanding of the basic laws of logic.

He misunderstands the basic laws of logic when he infers that, because an assumption used to build a model does not invariably hold in practice, the model results cannot be valid. This, in logical terms, says that the statement 'A implies B' is equivalent to the statement that 'not-A implies not-B'; clearly these statements are not equivalent — a fact which the author himself discovers when, in Section 11.4, he finds that it is possible to arrive at Black–Scholes results using different assumptions. This failure of logic is stated specifically in ¶20.2.3, but is also implicit in many other sections of the work.

The author also misunderstands the modelling process. He is critical of models where the assumptions used to derive results do not, in fact, hold; but this is exactly how much modelling is conducted; we make simplifying assumptions to construct a model; we get results; we test the results against the real world to assess whether the underlying simplification has invalidated the results. We try to develop better models which do not require the strong assumptions which we started with. The author appears to think that it has escaped the notice of finance professionals that some of the assumptions which they employ do not hold. This is not true — they know which assumptions are questionable, and they are, in general, intelligent enough to interpret the results appropriately.

Let me offer a pure actuarial analogy; for many decades life insurers have valued their liabilities using net premium reserves at a low, constant rate of interest. Nobody actually believed in the assumptions underlying this model, but we knew that, broadly, the results were accurate enough, although they required sensible interpretation.

Elsewhere in the paper the author tries to demolish well-developed theories by demonstrating how they may be badly applied; his Crusoe example of Section 2.6 does not invalidate the use of stochastic

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

models, it merely demonstrates the problems arising from using a poor statistical model. In Section 13 he tries to persuade us that there is no usefulness in game theory, because von Neumann & Morgenstern came up with a different strategy to Conan Doyle for Sherlock Holmes in the story *The Final Problem*; but von Neumann & Morgenstern were merely illustrating their ideas, not attempting a 'proof by example' (hardly something we would expect from one of the greatest mathematicians of the last 100 years); also, the author's objections are, in essence, based largely on a disagreement on the payoff matrix. This does not demonstrate that game theory itself is flawed. The author could have looked at any one of thousands of proven real world applications of game theory; why choose an untestable example from a work of fiction? It may be worth pointing out that an enormous quantity of development has taken place since von Neumann & Morgenstern laid the foundations of the Theory of Games in 1944, none apparently known to the author.

Another part of the paper which I found unpersuasive was the periodic table of ¶31.4. It seems to be a highly arbitrary, subjective matrix classification of 'theories' — though these do not appear to be well defined (I do not, for example, understand exactly what is included in the FE category). There is, of course, no reason why the author should not tell us his opinion of these 'theories' — but to call it a Periodic Table does not render it scientific or objective — two important features of the real periodic table. In fact, apart from being in table format, I can see no connection with the periodic table.

Some actuaries who dislike financial economics, perhaps because of a natural disinclination to accept techniques which are new and difficult, may find comfort in this paper; I suggest that this comfort will be short lived. The theory that the author finds so distasteful has permeated every major bank and investment house. This is because the finance professionals are delivering results, not because the banks are naive. It is important to understand that to ignore the everyday techniques of financial economics is to concede that actuaries are to be increasingly unimportant in financial circles. We should not accept this lightly.

Professor G. T. Pepper, C.B.E., F.I.A.: As an actuary who has followed an academic career in later years, I would have liked to have spoken at the discussion. I did not do so because the concerted attack on the paper came late in the session, there was too much that I wanted to say and time had run out. What follows should be read as if it were an oral contribution made at the meeting.

One of the first things that I noticed on reading the paper was that there were no acknowledgements. This suggested that the normal process at a university of circulating a paper to colleagues and benefiting from their comments might not have been followed. Whereas I wholeheartedly support the author when he challenges the simplifying assumptions underlying academic theory, he appears not to be in touch with recent developments. For example, there is now growing use of semi-variance (or, more generally, thresholds of tolerable loss) rather than variance, and modern analysis certainly allows for two classes of investor.

The features of the paper that worried me most were its hostile tone and the sweeping assertions that academic theory has nothing of value whatsoever to offer. I would have been in complete agreement if the author had confined himself to arguing that application of theory to practical problems can be very dangerous if the simplifying assumptions on which the theory is based are ignored. However, he does not do so, and the tone of the paper appears to jeopardise any chance of him communicating with academics.

I have deliberately not adopted such a hostile tone in my own work. My recent book, *Money, Credit and Asset Prices*, does not claim that the efficient market hypothesis is wrong; it merely states that a deduction from it is incorrect. In fact, I have changed my position since then. In 'The liquidity theory of asset prices', *The Actuary*, **5**, No. 11, August 1995, I argue merely that there is a special case, which many academics have not spotted, that is important in practice. I then advance the liquidity theory of asset prices, as developed by Maynard Keynes, Milton Friedman and James Tobin. This theory is not a rival to the efficient market hypothesis, but complementary. I show how a combination of the two theories explains much that the efficient market hypothesis cannot explain alone. I argue merely that the efficient market hypothesis is, by itself, an incomplete theory, which is something that should be easy for academics to accept. Even so I have run into resistance, so I have something in common with the author.

In ¶¶1.2.2 and 7.12 the author was kind enough to mention the model of the gilt-edged market that I developed in the early 1960s. At the time it was an adequate explanation of the market. Inflation then rose; so did interest rates; and gilt-edged stocks were issued with progressively higher coupons. The disparity in coupons became such that a two-dimensional model was no longer adequate and a three-dimensional one was needed. I almost took a sabbatical to try to develop one myself. I am very glad that I did not, because I would never have come up with an idea as original as the author's. It was too different, but, above all, it was too simple.

It is time that the full story of the author's gilt-edged model, described in Section 7, is put on record, and I hope to take steps to do this before long. At present I want to make three points which are particularly relevant to this discussion. The first is ironic. The Clarkson gilt-edged model was based on 'if you can get something for nothing, the market cannot be in equilibrium'. It was based on a simple version of the efficient market hypothesis, which the author is now attacking!

Secondly, some people will not be surprised to learn that the firm of stockbrokers that applied the author's model, mentioned in ¶7.15, was the one with which I used to be associated. I visited many clients to explain the new approach. Someone who was not technical could understand it quite easily, but most actuaries found it difficult. They could not forget about compound interest, redemption yields, differences between capital and income, and taxation. Their minds were programmed in the wrong direction. I will imply later that the same can happen to academics.

Thirdly, experience with the Clarkson model illustrated the efficient market hypothesis at work. As more and more people applied the new technique, the anomalies it disclosed became progressively smaller, until they were no longer worth exploiting. The market had again become 'efficient'. The author is right to argue, in ¶7.15, that some stockbrokers never understood the model and suffered as a result, but the efficient market hypothesis does not state that everyone has to adopt a new technique. It merely states that sufficient will do so for profit opportunities to disappear. This is what happened.

From what I have written so far, it will be realised that I had thought, and still think, that the author is at fault, particularly for the tone of the paper. Then came the concerted attack towards the end of the discussion. Two wrongs do not make a right.

In many disciplines there is a general tendency for academics to specialise; but to be a 'good specialist' you need to be a 'good generalist', in order to keep the specialisation in perspective. As the depth of knowledge becomes greater, people may be forced to specialise in progressively narrower fields. The result can be that they become very narrow minded and lose their perspective. Further, there is a tendency to build layer upon layer of complication on top of basic assumptions that are approximations. The result is an inverted pyramid on top of dubious foundations. As an intellectual exercise it may be very clever, but it has lost touch with reality. The academics would do much better to re-examine the foundations, secure them, and rebuild the pyramid. The difficulty is that they may have become so blinkered that they are incapable of doing so. Anyone querying the foundations is attacking their field of specialisation, which will become obsolete if the attack is successful. If this happens, the academics will become redundant if they are incapable of changing fields. Not surprisingly, they try to defend themselves.

I do not personally know the academics who made the concerted attack on the paper, (let me stress that I was listening to the discussion and not reading the written contributions that will appear in edited form in due course in *B.A.J.*), but those who spoke before Professor Wilkie sounded to me as if they had lost their perspective. They should, at least, have noticed that all of the experienced investment managers who had spoken before them had supported the author's assertion that the real world does not conform to their theories. They should have realised that there was something to learn.

The part of the paper that interested me most was the discussion about pattern recognition compared to logical thought. I think that the author is right that original thoughts often come from recognising patterns and are then developed using logic. The two mental processes are different. The author has a record of original thinking. He is currently groping in the dark to recognise patterns, in the hope that other original thoughts will follow. He should not be criticised for lack of logic in this part of his analysis. One sort of mind should respect the other.

Finally, ex-practitioners can be very helpful to academics, provided that there is communication between them. If I was a young academic, I would ruthlessly pick the brains of ex-practitioners and learn about the features of the real world that current theory cannot explain. I would concentrate on developing something new, and not on an endless extension of fashionable analysis.

Mr J. Plymen, F.I.A.:

FINANCIAL ECONOMICS — A SUPPLEMENTARY NOTE

"The market's pricing mechanism remains based to such a degree upon faulty and frequently irrational analytical processes that the price of a security only occasionally coincides with the intrinsic value around which it tends to fluctuate."

Security Analysis, by Graham & Dodd, 5th Edition, 1987

1. *'Modern' Portfolio Theory*

The description 'modern' is now a misnomer, as portfolio theory is still based on the work of Markowitz some 40 years ago. Unfortunately, the original assumptions regarding investment theory and risk made by Markowitz in 1952 are open to criticism in four areas:

(1) The expected return, whether based on historic data or investment analysis, is after specific risk, whereas Markowitz assumed that these returns were before risk.
(2) Markowitz's approach presupposes a perfect market which disregards the cheapness or dearness of individual stocks.
(3) The expected yield, the obvious criterion of performance, is ignored.
(4) The statistical base for MPT is purely historic, accepting the permanence of past short-term trends. There is no long-term element in the history.

It is not generally realised that the portfolios considered by Markowitz included bonds as well as common stocks. For the redeemable bonds, the major part of the risk lies in the prospect of disaster. In the event of a disaster, the bond or stock disappears from the market and becomes valueless. In these cases it is possible to obtain the performance before the risk by using the records of the surviving bonds rather than studying the performance of the whole bond market, including the failures. Only for bonds is it possible to obtain an indication of performance before risk. For all stocks that survive the period, any adverse influences in the past have been reflected in lower prices, so that the performance based on historic prices is after risk. Ignoring, for present purposes, the disaster risk which is only a tiny element of the whole, the performance record already allows for risk. Consequently, it is inappropriate to load the performance with the variance, thus imposing another layer of risk. In fact, mean-variance analysis, by charging up the risk against the performance twice, presents a completely misleading picture of the expected return.

Forty years ago investment analysis was in its infancy and computers were so slow and expensive that it was impossible to price shares sufficiently accurately to determine whether they were cheap or dear. Share price movements appeared to follow the random walk and the efficient market hypothesis was a tenable assumption. Operational research (OR) style quality control techniques are now used to study the past performance and to indicate whether the shares are cheap or dear in the short term. The standard deviation is quite irrelevant to the risk, merely determining the range over which prices can move.

For selecting securites, the level of the expected return after risk is clearly the sole criterion. Markowitz reckons that this criterion, applied to the whole portfolio at a given point of time, will result in putting the whole fund into a single share with no diversification. In practice, it does not work this way. Selection is only required when buying or selling a part of the portfolio. In any event, the expected yield is modified by the short/medium term cheapness or dearness, giving a degree of diversification.

The fourth weakness, the short-termism, is an inescapable feature of the Markowitz model. This means that, even if the first two disadvantages mentioned above are allowed for by scrapping the variance risk factor and by abandoning the efficient market, the model still lacks any guidance to long-term changes in market expectations.

At this stage Markowitz introduces the concept of the efficient portfolio. A portfolio is described

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

as efficient if there is no other alternative portfolio with either a higher average return and the same standard deviation or a lower standard deviation for the same return. In practice, the standard deviation is replaced by the variance as being more convenient. In effect the return and variance are traded against each other on a one-for-one basis. In view of the double counting of the risk, this efficient portfolio concept is meaningless. By selling off the highly performing shares too soon, it 'efficiently' reduces the performance. These highly performing shares are sold because they appear to have a high downside risk. This manoeuvre ignores the fact that these shares have an equal upside potential.

It must be borne in mind that Markowitz was writing around 1952, only 23 years after the United States common stock market nearly disappeared in a frightful slump. Consequently he was concerned about the downside risk. He shows charts of the prices for the previous 18 years of nine stocks, drawing attention to the size of the deviations below the mean as an indication of possible losses. The deviations above the mean are, of course, to the same order, so that the average overall historic profit and losses on either side of the mean balance out. By stressing the downside risk and ignoring the upside risk, he appears to be thinking in terms of bonds, which clearly carry a downside risk and no upside potential.

2. *Graham & Dodd*

In place of MPT, I prefer to go back to the portfolio theory of Graham & Dodd (1951).

> "The fundamental analyst believes that stock market prices fluctuate around the underlying investment value and that the two coincide occasionally. This belief rejects the idea that normally stocks are priced correctly in the market and even implies doubt that the market behaves rationally most of the time. Because stock prices move randomly with only a weak gravitational pull towards intrinsic value, the coincidence of price and value occurs infrequently, like Halley's comet. This means that the convergence time is long — perhaps three to five years."

In practice, the expected yield is derived from long-term investment analysis, making use of all relevant factors such as the asset value, the return on capital employed, the earnings/dividend cover, etc. The return calculated in this way is clearly after the intrinsic risk of the individual security. This return is, therefore, the sole criterion of success and needs no adjustment for 'risk'.

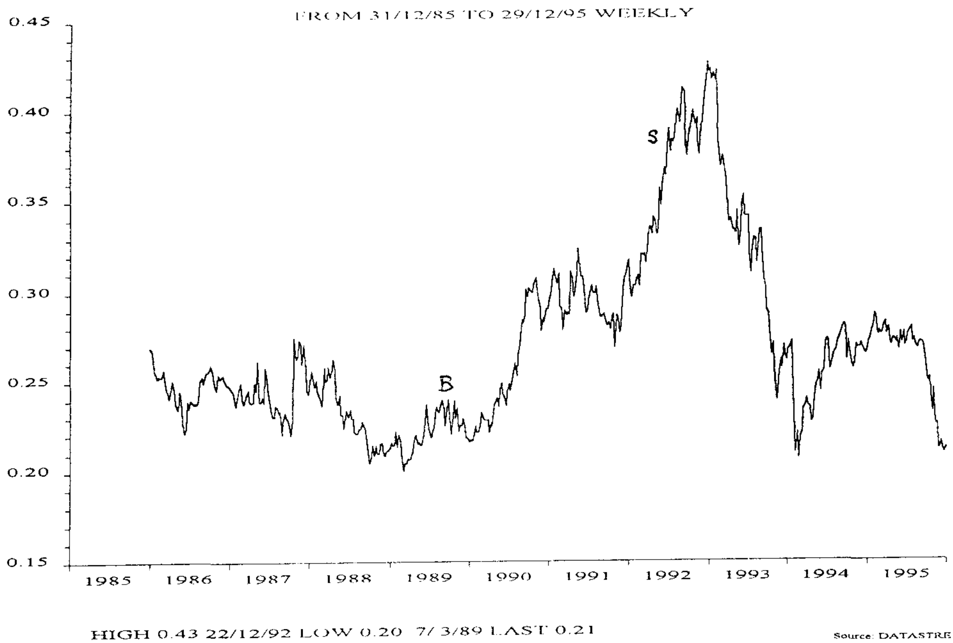
3. *Short Term / Long Term*

The Sainsbury share price, relative to the All-Share Index, is plotted and appears below. From 1986 to 1989 the long-term price index relative is tending slightly downwards. This is probably due to profit taking, as the share price had already risen fifteen times since 1979. This company closes the accounts by 31 March and reports the dividend in June. Every spring the price fluctuates as analysts speculate on the likely dividends and profits. Usually prices move from peak to trough and back again during the year, making two prices cycles within an average period of 6 months. The price gain from trough to peak varies around 18% to 22%. From 1990 onwards the long-term price trends are clearly changing, so that it becomes impossible to measure the size of the pure short-term changes.

The original Graham & Dodd principles involve policy moves that improve the expected long-term yield, but do not advocate any sort of trading between the lows and highs of the short-term fluctuations. I would point out, however, that such trading from the short-term price anomalies was not practicable earlier, because of the lack of statistics and the limitations of the early computers. Now that computer facilities, graph drawing facilities, etc. are so enormously improved, the prospect arises of possible trading around the short-term price fluctuations. If the swing from trough to peak is around 20%, it should be possible to cover costs and jobbers turns of 2% each way and to get into the share within 3% at the bottom and sell out within 3% at the top, and still leave a profit of 10% over 6 months, or 20% over the year if it can be done twice.

Clearly this short-term trading must be confined to periods when the long-term trend is more or less level. Between 1990 and 1994 the long-term trends have clearly taken charge, extending the peaks on the way up and the troughs on the way down. In these circumstances, the trading period must be increased from 6 to 12 months to 3 or 4 years (the convergence time of 3 to 5 years mentioned by Graham & Dodd).

SAINSBURY/ALL-SHARE INDEX



If a share is bought really cheap some degree of over-performance is axiomatic, as the price anomaly is corrected. As an example, the Sainsbury holdings of the OR Society Fund were bought in 1989 with a price/index relative around 0.24. Sales were made during the same half of 1992, with a relative factor of 0.38, just below the peak value of 0.42. The under-performance which started in December 1992 continued during most of 1993, followed by what chartists would call a 'head and shoulders' in 1994 and 1995. The price/index factor gained 58% over 3 years or some 16% p.a. as the price rating moved from being cheap both long and short term and being dear on both counts 3 years later. For this share the manager's skill rating is surely 100%, as the purchase was made under favourable terms in 1989, the shares were held while the price was doing well and the temptation to take a profit was resisted. Finally, the shares were sold just short of the peak relative value, anticipating the price collapse during 1993.

This operation is a prime example of the Graham & Dodd principle, whereby a share is purchased exceptionally cheap long term and is held until the anomaly is eliminated.

4. *Timing*

At this stage consideration must be given to the size and nature of the short-term fluctuations, as discussed already. The short-term variability has three components:

(1) the day to day fluctuations:
(2) the error term in the valuation model; and
(3) the fluctuations due to varying assumptions about approaching earnings and dividends.

The first two items appear in the form of 'grass' at the bottom of a radar screen, and have thus attracted this description. For all three items the dispersions from the mean must surely be completely random and completely independent. In practice, the majority of the price discrepancies will be in the

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

short-term category, with only perhaps 10% or so having a long-term element. Consequently, the overall errors from long and short-term discrepancies combined will closely follow a binomial distribution, which means that control lines at 1.57 times mean absolute deviation from the mean will cover most of the observations and will indicate the turning points for purchases and sales.

With the normal distribution there is clearly no absolute risk involvement. Downside risk, which is such a bug-bear to Markowitz, would be exactly balanced by upside potential.

At first sight, the return from a particular share appears to be dependent on just two items: the income; and the capital gains. Practical experience shows, however, that there is a most important third factor, i.e. the time element. The skilled analyst, after studying the accounts in considerable detail and after close liaison with the managers, may spot, perhaps, a long-term disaster well before the market appreciates the position. I can remember selling a large line of Vehicle and General shares for an important institutional client. For the next 6 months the price went up steadily to a level some 80% higher. Shortly after that the company became insolvent and the shares were worthless. In this case my adverse view was based on a high level of creative accountancy which was spotted by me, but not by any other analyst. Superficial analysts, influenced by the satisfactory profits and dividends, ranked the shares as short-term cheap and pushed the price up accordingly. This shows that an important sale decision should never be taken without studying the short-term price history and outlook.

Stockbrokers, when recommending a sale of a particular share, are very tempted to accompany this recommendation with a purchase suggestion, probably in the same industry. Often a switch transaction, which involves a close comparison between features of each stock, appears easier to justify than the mere sale and re-investment in the index fund. For static funds like investment trusts, the purchase recommendation may need to be accompanied by a sale transaction to provide the funds.

In practice, a well-managed active institution should not have to balance off purchases and sales for monetary reasons. The institution should, as a matter of policy, keep a substantial float, preferably invested in an index fund, which is drawn upon to finance purchases and to absorb sales. This means that purchases and sales can stand on their own feet, so to speak, and can be made at the most advantageous time for each share. With a switch, it would be a miracle if the purchase was sufficiently cheap and the sale sufficiently dear to justify the individual transactions. Within a particular industry, the sale may be justified because of the deteriorating conditions for that trade. In such circumstances it is clearly better to make the sale, reinvest in the index fund and await a better opportunity for buying the other share in the same group.

5. *The Dynamic Portfolio*

In place of MPT, I suggest a new form of portfolio theory designed to give a maximum performance by an active policy of trading, exploiting short and long-term anomalies. I assume that the manager has available a highly sophisticated equity market model, which can given an accurate pricing to an individual share and will show, at any time, whether the share is cheap or dear in both the long and the short term. For this purpose the Clarkson market model, with perhaps certain updating, should be satisfactory. I recommend a simple procedure, that is to buy shares that are exceptionally cheap and are, consequently, liable to over-perform in the future. The purchases are closely monitored against the index, and the shares are held so long as the price is over-running the index. In due course the price anomaly will be corrected, and the share becomes fairly priced and will be sold. In effect, this means that every share in the portfolio is over-performing the index; shares are only held over the favourable phase when they are over-performing the index. Every holding is regarded as a short-term operation, going into the share when it is cheap and running it until a fair price is obtained.

6. *Dynamic Indexation*

It is important, when applying such a policy, to make sure that sales are not made too soon. So long as the share is cheap long term, it should be held as a further potential for profit. The judgement of the sales at the turning points is assisted by reference to the sophisticated market model, with its cheapness or dearness assessment. The procedure for operating this 'cheap only' investment programme is as follows.

https://doi.org/10.1017/S1357321700004797 Published online by Cambridge University Press

Start off with an index fund worth £1,000m. Whenever a suitable share is available at an advantageous price, say £5m is taken from the index fund and put in the cheap share. This holding will then be run for sufficiently long to realise its full potential. It may be cheap short term, in which case the transaction can be reversed within a year or two. It may, however, be cheap long term, in which case the run off of the profitable operation may take 3 or 4 years. (See note about performance of Sainsbury shares over 1988 to 1993.) Gradually, as the cheap shares become available, further holdings of £5m in each can be taken up. When each holding has done its job and realised a profit, the share is sold and the money put back into the index fund. Ultimately there may be, say, 100 shares that are still in the cheap phase and are included in the dynamic portfolio as opposed to the index fund of £500m remaining.

The author subsequently wrote: My suggested framework for downside risk, which is, in effect, the mathematical cornerstone of my paper, was subjected to two very severe attacks; Mr Smith and Dr Bowie suggested that my theoretical framework was nothing more than a repetition of concepts first developed in 1977 by Bawa & Lindenberg in their paper 'Capital Market Equilibrium in a Mean-Lower Partial Moment Framework' (*Journal of Financial Economics*, **5**, 189–200), and Mr Smith and Dr Cairns attempted to show that my key numerical example in Section 11.5 was invalid. Although I responded in a preliminary fashion to each of these two areas in my verbal reply at the meeting, I believe that the written record should contain a fuller response to these severe criticisms which, if justified, would, of course, seriously undermine both the theoretical and the practical relevance of the entire paper.

The very first sentence of the Bawa & Lindenberg paper is:

> "Portfolio selection is a decision problem involving a choice among elements of a set of known probability distributions of returns."

I disagree; the crucial point about real world investment decisions is that the underlying probability distributions are *not* known with certainty, and the observations by R. L. Gwilt, that I cited in my verbal reply, are very relevant. The incompleteness principles that I develop in Section 22 are put forward as a more plausible starting point.

Bawa & Lindenberg then discuss 'admissible portfolios' on the assumption that investors have von Neumann & Morgenstern utility functions; I do not regard this as a reasonable assumption, and hence develop my approach along quite different lines.

Bawa & Lindenberg regard the optimal portfolio as the one which maximises the expected utility of return; I reject the expected utility approach as being theoretically unsound.

Bawa & Lindenberg derive their capital market equilibrium framework on the assumption that all investors have homogeneous expectations about future returns; I regard this assumption as grossly unrealistic, and develop, instead, my different stereotypes of investors, and show, in Section 6.9, how it is possible for three investors with the same information set to come to totally different conclusions as regards expected returns.

Most importantly, Bawa & Lindenberg do not incorporate the explicit maximum level of acceptable risk that is an essential feature of my formulations of optimal and rational behaviour. For all these reasons, the Bawa & Lindenberg framework for capital market equilibrium does not bear the slightest resemblance to the alternative framework that I have put forward for discussion.

Several points need to be stressed in relation to the criticisms that Mr Smith and Dr Cairns made of my numerical example of risk in Section 11.5. Firstly, the two distributions intersect at such a strongly negative value of return, and at such an infinitesimally small value of probability density function, that, for all practical purposes, it is indeed reasonable to say that 'one distribution lies to the right of the other'. Accordingly, it would be unrealistic in the extreme to use a risk threshold lower than the return represented by this point of intersection. Secondly, a very convenient way of dealing with a normal distribution in practical numerical work, such as that arising in asset-liability modelling or in capital projects, is to replace it by a much more tractable binomial distribution with the same mean and variance. Clearly the accuracy of the approximation improves as the number of points in the binomial distribution increases, and I have found that the 7-point approximation, where the probability values, after multiplying by 64, are 1, 6, 15, 20, 15, 6 and 1, usually gives a good balance between accuracy and ease of computation. If we take the distribution N(5,2) in my example, the crude 3-point approximation involves the values 3, 5 and 7 with probabilities 0.25, 0.5 and 0.25 respectively. The downside risk value for a risk threshold of 6 and a risk weighting parameter of 2 is then:

$$0.25 \times 3^2 + 0.5 \times 1^2 = 2.75$$

which differs by less than one half of one percent from the more accurate value of 2.739 that I obtain using the 7-point approximation. Thirdly, neither Mr Smith nor Dr Cairns were able to find the underlying reason for what they regarded as a paradoxical conclusion, namely my assertion that utility theory leads to a completely different result from my downside risk approach. The reason for the conflicting results is that, as I show in Section 6.6 of Clarkson (1990), my downside risk approach leads to an implied local utility curve which is linear above the risk threshold and concave downwards below it. The utility curve in the example, however, is exponential in nature, and thus has a shape which is impossible in my framework.

The opener gave a commendably succinct description of the two competing paradigms, namely 'financial economics' as characterised by three broad features, and my 'proposed paradigm' as characterised by six quite different features. While admitting to a bias in favour of my 'proposed paradigm', on the basis of his own actuarial and investment experience, he also referred to certain evidence which he thought pointed the other way, and, in particular, he referred to statistical studies showing that investment managers of, for example, unit trusts find it very difficult to outperform consistently. This statistical evidence is, I believe, fully consistent with my 'proposed paradigm', and two particular instances of market behaviour can be used to illustrate my arguments.

The switch out of engineering shares, referred to in Section 8.8, could be described as an instance of what I call 'intelligent behaviour', in that a major fundamental judgement, which subsequently proved to be correct, suggested that the profit potential was significant. However, since most other investors at that time came to a different (and, with hindsight, erroneous) conclusion, it took around two years for the switch to come into profit. Similarly, a general switch out of equities into cash in the first half of 1987, on the basis that the gilt/equity yield ratio had become unsustainably high, would have shown what many commentators would have interpreted as worryingly bad relative performance over the twelve months to 30 September 1987, the last end-quarter date before the strategic switch would have been fully vindicated by the dramatic collapse of equity markets worldwide in October 1987. On intuitive grounds, it seems reasonable to say that both these switches were instances of what the opener called 'more skilful' behaviour. However, as I stress in ¶28.1.2, it is exceptionally difficult in practice to predict when 'the market', which for a time can be dominated by 'unintelligent' behaviour, will move into line with a correctly anticipated major change in the fundamentals, and this is why I introduce a specific 'observe the behaviour of others' component into my formulations of optimal and rational behaviour.

This suggests that a more meaningful test for the existence or otherwise of investment managers who are significantly 'more skilful' than the average would be to examine rolling two-year periods rather than supposedly time-independent quarters or calendar years; this would, I believe, highlight the existence of superior performance that could not be dismissed as mere random chance. In this connection, it is interesting to note that a forthcoming paper by Professor Gruber in the *Journal of Finance* will use more sophisticated performance tests than those previously employed to show that the average performance of U.S. stock mutual funds has been much better than shown by earlier studies; these results could re-open the whole debate as to whether 'active' investment managers can achieve consistently superior performance, and thereby justify their fees.

Dr Cairns complained that there were several instances where I took the existence of a single dataset consistent with my ideas as proof that I was right and that financial economists were wrong. However, as Kuhn, Popper and other eminent philosophers of science have shown, scientific progress often takes place through the discovery of a completely new paradigm which can resolve previously detected areas of anomalous behaviour rather than through the continual refinement of the old paradigm. Accordingly, the existence of even a few relevant instances of a new paradigm being able to explain real world behaviour that is anomalous in the context of the old paradigm must surely be an indication that followers of the old paradigm have a scientific and professional duty to the community at large to investigate, in a balanced manner, whether, in the interests of society as a whole, there is a sufficiently strong case for the new paradigm to be adopted as the generally accepted scientific approach.

In his very extensive criticism of my non-linear inflation model, which is perhaps the most 'actuarial' part of my entire paper, Dr Cairns suggests, not only that I have dispensed with the use of mathematics and statistics, but also that I am completely irrational in arguing in favour of my model over other models such as the Wilkie model. The reality, I believe, is quite different. In 1981 and 1982, my colleague, Dr Namdar Mossaheb, and I carried out very extensive tests of various ARIMA models of the Box–Jenkins type to see if we could improve upon the already very good predictive power of Mean Absolute Deviation techniques in the analysis of gilt-edged prices. We carried out a large number of statistical tests along very similar lines to those subsequently published in Wilkie (1984), and we concluded that four apparently quite different times series all gave an equally satisfactory representation of the data — ARIMA 400, ARIMA 410, ARIMA 012 and ARIMA 111. Although this approach appeared to hold out considerable promise, there was evidence that the price behaviour was markedly non-linear in nature, with the result that the residual variance was so high that the model had virtually no forecasting ability. When the ARIMA-based Wilkie model was first discussed at the Faculty in 1984, I described our attempted application of ARIMA techniques to gilt-edged securities, observed that an inflation series could be expected to exhibit even more pronounced non-linearity, and suggested, as a tentative alternative inflation model, a formulation virtually identical to that set out in ¶12.7.1. Having remarked, at the end of my contribution to the 1984 discussion, that I was confident that an improved inflation model could be built fairly easily, I carried out further mathematical work on my tentative model a few years later, and wrote up this research in Clarkson (1991), which was presented at the Brighton AFIR Colloquium. As I state in ¶12.10.1 of the present paper, my main criticism of the Wilkie inflation model is that it does not generate high values that are closely bunched over time, and accordingly the crucial statistical test of my non-linear model is whether the normal variate percentile points are more random over time than is the case for the Wilkie model. Not only do I conclude, in ¶12.10.2, that, in the case of my model, the variates are much closer to being random, but I also show, in ¶12.13.7, that Engle (1984) applied precisely the same statistical considerations before concluding that his ARCH model represented an improvement over linear formulations. This question of time-dependence is particularly relevant in the light of the background to the Wilkie model, which arose out of the work of the Maturity Guarantees Working Party, which, in turn, was set up in response to a severe financial crisis in the 1970s when, for several successive years, inflation rates and interest rates were vastly in excess of values that might have been expected from any linear model fitted to past experience. Against the background of the mathematical work and statistical investigations outlined above, I regard Dr Cairns' criticisms in relation to my non-linear inflation model as unjustified.

Although Dr Macdonald criticises me for abandoning the precision of mathematics, my belief that an emphasis on principles rather than on abstract mathematical theorising is likely to be the best way forward is fully consistent with views expressed by Professor Maurice Allais in his Nobel Lecture of 8 December 1988:

> "The use of even the most sophisticated forms of mathematics can never be considered as a guarantee of quality. Genuine progress never consists in a purely formal exposition, but always in the discovery of the guiding principles which underlie any proof."

Furthermore, there are two areas, in particular, where I disagree with Dr Macdonald's criticisms of the logic of the derivation of my alternative approach.

The first area relates to my Robinson Crusoe analogy. Although Dr Macdonald attempts to invalidate my analysis, I believe that the essential themes remain intact — risk depends on the training and prior experience of the individual, utility theory has no relevance, and stochastic models, in view of their generally impenetrable 'black box' nature, are often unhelpful in real world applications that involve a very high level of risk.