The Coming Revolution in the Theory of Finance

Robert S Clarkson
"Cherrybank"
Manse Brae
DALSERF
Lanarkshire
Scotland
ML9 3BN

Telephone: 0698 882451
Facsimile: 041 204 1584

Summary

The aim of this paper is to highlight the unsound nature of many of the simplifying assumptions that are inherent in the currently accepted theory of finance, and thus identify which of these assumptions have to be abandoned before the widening gulf between financial theory and financial practice can be bridged. The empirical evidence suggests that the real financial world is dynamic and non-linear in nature, whereas the current theory is formulated within essentially static and linear frameworks such as the Capital Asset Pricing Model and the Arbitrage Pricing Model. Furthermore, it is shown that the use of historic variability of returns as a measure of risk is highly unsatisfactory, particularly when testing for market efficiency.

Révolution imminente de la théorie financière

Robert S. Clarkson
« Cherrybank »
Manse Brae
DALSERF
Lanarkshire
Ecosse
ML9 3BN

Téléphone : (44) 0698 882451
Télécopie : (44) 041 204 1584

Résumé

La présente étude vise à mettre en évidence la nature erronée d'un grand nombre d'hypothèses simplificatrices acceptées dans la théorie financière actuelle et d'identifier celles de ces hypothèses qu'il convient d'abandonner afin de réduire l'écart grandissant qui sépare la théorie de la pratique quotidienne. L'expérience pratique indique que le monde financier est de nature dynamique et non linéaire. Par contre, la théorie financière en son état actuel est formulée au sein de systèmes essentiellement statiques et linéaires, tels que le modèle d'évaluation des actifs financiers (MEDAF) et le modèle d'évaluation selon l'arbitrage. Il est prouvé par ailleurs que le concept de la variabilité historique des rendements ne permet pas d'évaluer le risque de manière satisfaisante, en particulier lorsque l'on vise à déterminer l'efficience des marchés.

The Coming Revolution in the Theory of Finance

Introduction

Over recent years there has been considerable divergence between the development of methods for practical financial management and the teachings of the theory of finance. Three specific divergences between theory and practice are of particular significance. Firstly, there is a growing awareness that linear models based on the expected utility maxim may not give an adequate representation of the real financial world. Secondly, the Sharpe Diagonal Model, with its emphasis on beta rather than alpha as the main component of investment return, implicitly assumes that capital markets are highly efficient in discounting all relevant information, whereas there is an increasing body of evidence to suggest that exploitable pricing inefficiences are far more common than was previously thought to be the case. Thirdly, many practical methods for assessing risk focus on simulations of unsatisfactory outcomes on the downside rather than using the symmetric measure of variance over the entire range of possible returns.

These specific examples of the widening gulf between financial theory and financial practice suggest that we should re-examine the following three key postulates that underly most practical applications of the theory of finance:

1. Investors behave "rationally" in that they are risk-averse and their actions conform to the expected utility maxim.

2. The Sharpe Diagonal Model is an adequate representation of the price formation process within capital markets.

3. Variance of return is a satisfactory proxy for risk.

These postulates were necessary assumptions that had to be made to obtain practical frameworks such as the Capital Asset Pricing Model, but the literature contains no direct empirical tests of whether or not these postulates are satisfactory. In the physical sciences, however, new and better theories have often resulted from experimental evidence that the predictions of earlier theories were not in accord with the real world. One of the classic examples was the Michelson-Morley experiment in 1887 which gave unexpected results regarding the speed of light; this led to the theory of relativity.

The aim of this paper is to repeat this experimental approach in the financial world by testing in turn each one of the three postulates described above. In each of the three cases an alternative hypothesis, radically different from the teachings of the theory of finance, is found to be consistent with the empirical evidence. All three of the postulates listed above must therefore be abandoned before the widening gap between financial theory and financial practice can be bridged.

The Expected Utility Maxim

Utility can be regarded as a conceptual device introduced to reflect the fact that human beings take risk into account when assessing future uncertain events rather than acting solely on the basis of expected value in probabilistic terms. Rational behaviour is then taken to be equivalent to maximising the expected value of utility, and this is generally referred to as the expected utility maxim. Markowitz (1959) discusses this crucial aspect of human behaviour in meticulous detail and comments as follows on the contradictions which arise:

> "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 towards irrationality, even in his most reflective moments."

After discussing three examples of apparently inconsistent behaviour, one of his own quoted in Alchian (1953) and two quoted by Allais (1953), Markowitz draws the conclusion that the individuals choosing the "wrong" alternative acted irrationally. The second of these three examples is, I believe, by far the most important. Let us now analyse it in detail to see if we can find an alternative explanation for human behaviour under conditions of risk and uncertainty. Translated into sterling, the example is as follows:

Allais asked subjects to consider the following two alternatives:

Alternative A:	receive £1,000,000 with certainty
Alternative B:	receive £5,000,000 with probability 0.1
	receive £1,000,000 with probability 0.89
	receive nothing with probability 0.01

He found that his subjects preferred alternative A to alternative B. He then asked them to consider the following:

Situation C:	receive £1,000,000 with probability 0.11
	receive nothing with probability 0.89
Situation D:	receive £5,000,000 with probability 0.10
	receive nothing with probability 0.90

He found that his subjects preferred D to C.

The situation can now be analysed by letting U(0), U(1) and U(5) be the utilities of receiving nothing, £1m and £5m respectively. The preference of A over B implies:

U(1) is greater than 0.1 U(5) + 0.89 U(1) + 0.01 U(0).

Adding 0.89 U(0) - 0.89 U(1) to both sides gives:

0.11 U(1) + 0.89 U(0) is greater than 0.1 U(5) + 0.9 U(0),

and hence an individual who follows the expected utility maxim and prefers A to B must also prefer C to D. Allais concludes that the expected utility maxim must be unreasonable, whereas Markowitz concludes that the expected utility maxim is indeed the criterion of rational behaviour but that Allais' subjects acted irrationally.

I not only agree with Allais but can also explain the flaw in Markowitz' logic very easily - expressions such as "the utility of £1m" have no absolute meaning but depend on the overall financial situation of the individual in question. Since the overall financial situation implied by the existence of alternatives A and B is radically different from that implied by the existence of alternatives C and D , we cannot assume that each of U(0), U(1) and U(5) has the same value in both situations.

To obtain an alternative hypothesis that can explain the actions of Allais' subjects, we need to assess risk relative to the "current wealth" of the individual. The introduction of this reference point against which risk is to be measured is a key principle of the alternative measure of risk that I describe in Clarkson (1989) and Clarkson (1990), but for the moment a "general reasoning" approach rather than an analytical and axiomatic approach is perfectly satisfactory.

The £1m available with certainty under alternative A can be assumed to be vastly in excess of the previous wealth of the individual, and now forms the benchmark against which alternative B has to be judged. Alternative B introduces a small probability of receiving £5m and also a minute, but non-zero, probability of receiving nothing. He is likely to regard £1m and £5m as similar "infinities" in relation to his previous wealth, and he will certainly not regard £5m as being five times as desirable as £1m; £1m will for all practical purposes satisfy all his desires as regards quality of life in terms of money. The minute probability of receiving nothing, however, will prey on his mind. He would never forgive himself for the rest of his life if he turned down the certainty of riches beyond his wildest dreams only for the one chance in a hundred to ruin what would otherwise have been a utopian existence.

In situations C and D, however, there is only about one chance in ten of a vast increase in wealth, and he knows that he is likely to receive nothing. He will assess the slim chance of riches under C and D against the benchmark of his current wealth rather than against £1m, and he will see himself as having nothing to lose by going for £5m rather than £1m; the 90% chance of losing when going for £5m is for all practical purposes the same as the 89% chance of losing when going for £1m.

In Clarkson (1990) the parallel analytic approach is used to obtain solutions to the St Petersburg Paradox which depend on the current wealth of the individual. The fatal flaw in the expected utility maxim is the assumption that an individual assesses all financial situations in terms of a single utility curve. The fact that risk has to be assessed relative to a benchmark position which will vary with the overall financial situation means that this assumption is untenable.

It is interesting to note that when Sharpe (1970) investigates analytically the implications of an investor using a fixed utility curve to assess portfolios in terms of expected return and standard deviation of return he arrives at a conclusion that is virtually identical to mine:

> "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."

The Sharpe Diagonal Model

While Markowitz' solution to the general portfolio selection problem as set out in Markowitz (1952) could be expressed elegantly in terms of matrices whose elements were covariances of return between various securities, there was initially no practical way in which the numerical values of these elements could be estimated. The breakthrough came when Sharpe (1963) suggested his "diagonal model", which in essence assumed that the future price of a security depended on three items, its "alpha", the market return through its "beta", and a random error term.

The increasing importance attached to the "Efficient Market Hypothesis" resulted in "alpha" being of far lesser importance within Modern Portfolio Theory than "beta", as can be seen for instance from influential textbooks such as Rudd & Clasing (1982), where examples of "excess alpha" (i.e. exploitable inefficiencies on a risk-adjusted basis) are clearly seen as being the exception rather than the rule. However, the literature does not contain any empirical tests of whether "beta" does indeed dominate "alpha" in the price formation process within capital markets.

The experiments that I carried out a number of years ago relate to the UK gilts market, on the basis that estimation errors regarding returns to maturity are non-existent (except to a very slight extent where the redemption date is variable at the option of the Government) and that dealing expenses are very low. If "beta" dominates "alpha" in equity markets, then an even stronger dominance should occur in the gilt market, and the price movements of a stock should, over the medium term, be able to be represented very accurately in terms of the "beta" and the market return.

The results showed that the expected linearity, with a high "beta" for very volatile low coupon stocks and a low "beta" for less volative high coupon stocks, was not present. There was, however, not only a very significant random error element but also a most intriguing "alpha" element - a clockwise loop (on a rise in the market followed by a fall) for low coupon stocks, and an anti-clockwise loop for high coupon stocks when the reciprocal of the running yield was plotted against the market level as measured by the par yield curve value in my non-linear gilts model as described in Clarkson (1978).

I believe that these results are highly significant for two reasons. Firstly, they suggest that "beta" will be even less important in equity markets, where in addition to pricing inefficiencies which correspond to high random errors the deterministic cyclicality of "alpha" has an obvious parallel in "sector rotation", i.e. investor sentiment moving from one economic theme to another in anticipation of possible future economic scenarios. Secondly, this deterministic cyclical behaviour throws considerable doubt on an apparently innocuous assumption of the Capital Asset Pricing Model, namely that the capital market is in equilibrium.

Variance of Return as a Proxy for Risk

A key postulate of the theory of finance is that investors are risk-averse, where risk is measured in terms of some statistical measure of short term price variability. Moreover, a symmetric measure, such as the variance of return or its equivalent the standard deviation, is usually used in preference to the more intuitively appealing types of downside measure such as semi-variance. Text books on Modern Portfolio Theory often state that risk is equivalent to variability of return without giving any justification whatsoever. For example, Rudd & Clasing (1982) describe risk as follows in the glossary:

> "The uncertainty of investment outcomes. Technically, the term risk is used to define all uncertainty about the mean outcome,

> including both upside and downside possibilities. Thus, in contrast to the layperson, who would think of the downside outcome as risk and of the upside outcome as potential, a measure of total variability in both directions is typically used to summarize risk."

Given my rejection of the expected utility maxim, which attempts to combine expected return and risk (as measured by variability of return) into one composite measure of attractiveness, it is highly desirable to formulate and test an alternative hypothesis regarding the nature of short term price variability.

Consider first two illuminating comments about the non-stationary nature of price variability. Mandelbrot (1963) begins the final section of his highly controversial paper as follows:

> "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 of prices in periods of tranquillity seems 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."

Fama (1965) also shows that large daily price changes tend to be followed by large changes, but of unpredictable sign. In his major review work on stockmarket efficiency, Fama (1970) provides a possible mechanism for this phenomenon:

> "This suggests that important information cannot be completely evaluated immediately."

My alternative hypothesis is a generalisation of these observations, namely that short term price variability is a function of the frequency with which important new information arrives that has not already been discounted in consensus forecasts.

Practical experience of equity investment suggests that the frequency with which important new information has not been discounted in "consensus" forecasts is very much a function of the type of company. At one extreme we have utility companies, where in general there are government-controlled limits on the profits that can be earned. The mechanism for this control may be through the rate of return on capital employed, or through a cruder basis such as limiting the maximum price rise for the commodity in question to X% less than the rise in an index of inflation. Since there are relatively few "surprises" regarding profits for this type of company, the short term price variability is much lower than the market average.

High technology companies are at the other end of the spectrum; the potential rate of return on capital employed can be very high, but new product breakthroughs, produce obsolescence, price cutting by competitors, take-over speculation etc. etc. result in frequent announcements of important new information that cannot possibly be forecast with any degree of precision. Keynes (1936) describes very vividly the impossibility of forecasting the profitability of specialist companies:

> "The outstanding fact is the extreme precariousness of the basis of knowledge on which our estimates of prospective yield have to be made. 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."

As a result of the frequent announcements of quite unpredictable new information, short term price variability for high technology companies is much higher than the market average.

We now need an appropriate source of risk-return data of the type used in the Capital Asset Pricing Model to test my alternative hypothesis. One of the most successful scientific discoveries in the physical world was Kepler's conclusion that the orbits of planets around the sun were elliptical rather than, as was previously thought to be the case, circular.

Kepler's results depended in large measure on the very detailed observations of Tycho Brahe, who had produced by far the most accurate set of astronomical data available at that time. The obvious parallel in the case of the theory of finance is the very comprehensive risk-return data compiled by Jensen (1968) in connection with his "strong level" efficiency tests on US mutual funds. His sample of 115 funds includes 4 science or high technology funds and one utilities fund.

Jensen uses the framework of the Capital Asset Pricing Model to test for "strong level" efficiency. The "nul hypothesis" for my experiment is therefore that there is no evidence in Jensen's data to suggest that there is any correlation between the "alpha" (i.e. the risk-adjusted excess return) as calculated by Jensen and the nature of the mutual fund. If my alternative hypothesis is correct, the very high short term price variability in the case of high technology funds will tend to result in a negative "alpha", whereas the very low short term price variability in the case of the one utilities fund within the sample is likely to cause it to be the fund with the highest "alpha", regardless of its actual return.

The results provide such clear evidence in favour of my alternative hypothesis that further statistical examination is unnecessary. Of the 115 funds in the sample the rankings of the 4 science or high technology funds were 76, 89, 110 and 115; all in the bottom half, and including the very bottom fund. The one utilities fund was in first position by an implausibly high margin; its risk-adjusted excess return was more than two and a half times that of the fund in second place.

Three inferences can be drawn from this investigation:

(i) It is unsound in general to equate variability of return to investment risk.

(ii) In the light of (i), the Capital Asset Pricing Model cannot in general be used as a measurement framework for risk-adjusted excess return.

(iii) In the light of (ii), Jensen's highly influential conclusions regarding the existence of "strong level" efficiency are unsound.

Conclusions

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.

References

Alchian, A.A. (1953). The Meaning of Utility Measurement, American Economic Review, Vol. 63, No. 1., (March, 1953), p.26.

Allais, M. (1953). Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. Econometrica, Vol. 21, No. 4 (October, 1953), p. 503.

Clarkson, R.S. (1978). A Mathematical Model for the Gilt-Edged Market. Transactions of the Faculty of Actuaries, Vol. 36, p. 85.

Clarkson, R.S. (1989). The Measurement of Investment Risk. Transactions of the Faculty of Actuaries, Vol. 41, p. 677.

Clarkson, R.S. (1990). The Assessment of Financial Risk. Proceedings of the 1st AFIR International Colloquium, Vol. 2, p. 171.

Fama, E.F. (1965). The Behavior of Stock Market Prices. Journal of Business, 38 (January, 1965), p. 34.

Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25, No. 2 (May, 1970), p. 383.

Jensen, M. (1968). The Performance of Mutual Funds in the Period 1945-64. Journal of Finance, 23, p. 389.

Keynes, J.M. (1936). The General Theory of Employment, Interest and Money. Macmillan.

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. Journal of Business, 36, No. 4 (October, 1963), p. 394.

Markowitz, H.M. (1952). Portfolio Selection. The Journal of Finance, 7, No. 1 (March, 1952), p. 77.

Markowitz, H.M. (1959). Portfolio Selection: Efficient Diversification of Investments. Second Edition (1991). Blackwell, Oxford.

Rudd, A. and Clasing, H.K. (1982). Modern Portfolio Theory: The Principles of Investment Management. Dow Jones-Irwin.

Sharpe, W.F. (1963). A Simplified Model for Portfolio Analysis. Management Science, 9, No. 2 (January, 1963), p. 277.

Sharpe, W.F. (1970). Portfolio Theory and Capital Markets. McGraw-Hill, New York.