COWLES COMMISSION DISCUSSION PAPER: ECONOMICS: NO. 2005

NOTE: Cowles Commission Discussion Papers are preliminary materials circulated privately to stimulate private discussion and are not ready for critical comment or appraisal in publications. References in publications to Discussion Papers (other than mere acknowledgment by a writer that he has had access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

Statistical Problems and Computational Programs Suggested by the Theory of Investment Behavior

by

Harry Markowitz

January 23, 1951

Introduction.

Cowles Commission Discussion Papers Economics 278, 294 and this paper deal with the problems of deriving behavior equations to be fitted to and tested by data concerning Investment Companies. We do not seek equations to be "believed in" a-priori. Rather we seek to derive equations from interesting, plausible or traditional economic hypothesis. If such equations are fitted, tested critically and rejected then an interesting economic hypothesis has been rejected. If they are tested critically and not rejected, then the underlying economic hypothesis, in some sense, gains in plausibility. In either case the behavior equations have served one purpose, and may serve more.

The writer is more interested in behavior under uncertainty and financial behavior in general, than in investment company behavior in particular. The open-end investment company, however, seems ideal as a place to start in a study of these matters. Its institutional simplicity gives hope of deriving verifiable, computational manageable behavior equations from plausible economic hypothesis.(1)

(1) The typical open-end investment company sells its stock to the public and uses the funds to buy securities which are nationally traded in the open market. The open-end investment company stands ready to buy back its stock at "asset value" (i.e., at a price equal to the current value of the securities held by the company divided by the number of shares of stock the company has outstanding). The investment company typically has no fixed debts. All interest and dividends which the

The deriviation, fitting and testing of behavior equations, require the joint efforts of the economist, mathematician, statistician and computer. The present paper is based on this need for a division of labor. In Part A we present (and explain the source of) some problems of identification and estimation. Those problems have been met in the pursuit of investment company behavior equations. In Part B, we present a hypothesis which is ready to be fitted to and tested by the relevant data.

Part A

In Cowles Commission Discussion Paper Economics 294 we assumed that the investigator could specify (a) which moments entered the utility function; (b) the forms of belief formation functions (b.f.f's) $\mu_i = \mu_i(Z_1,\ldots,Z_s)$, $\sigma_{ij} = \sigma_{ij}(Z_1,\ldots,Z_s)$ etc.; we also assumed (c) that these b.f.f's (belief formation functions) are linear in any unknown parameters. From these assumptions we derived behavior equations.

We will argue here that if we assume that random elements enter these equations via the usual random "shocks", then the equations are identifiable. Even in this shock model, problems of estimation arise. But the shock model is inadequate. Models derived from more reasonable assumptions as to how random variables enter, will be presented.

For purposes of exposition let us consider the equations which follow if utility (U) is assumed to depend on expected value (E) and variance (V). The extension of our discussion, to the general case, will be given later.

In Discussion Paper Economics 294 we saw that the hypothesis that $U = U(E, V)$ implies the equations

---

(footnote (1) continued) company receives on the securities it holds is paid out, after expenses have been deducted, to the holders of the investment companies'. stock. (This is to take advantage of tax provisions). The compensation to the management is proportional to the total assets of the company. There is almost no possibility of failure. There is little problem of liquidity. Almost no deviation of interest between management and stockholders.

$$(1) \qquad \sum_{j=1}^{n} X_j \left\{ (\mu_2 - \mu_1)(\sigma_{ji} - \sigma_{1i}) - (\mu_j - \mu_1)(\sigma_{2i} - \sigma_{1i}) \right\} = 0$$

$$(\Sigma X_j = 1) \qquad\qquad i = 3, \ldots, n$$

If the b.f.f's $\mu_i = \mu_i(Z_1, \ldots, Z_s)$, $\sigma_{ij} = \sigma_{ij}(Z_1, \ldots, Z_s)$ are linear in unknown parameters α and β, respectively, then the behavior equations are linear in unknown parameters $\delta = \alpha\beta$.

Suppose we add "shock terms" u_i to equations (1); where $\mathcal{E} u_i = 0$ for all i, $\mathcal{E} u_i Z_j = 0$ for all ij. The resulting model is typically identifiable.(2)

Problems of estimation arise because the variables do not enter linearly as is usually assumed. Each equation involves a sum of terms of the form $\delta_{k\gamma} X_i Z_k Z_\gamma$. Each term is the product of exogenous and endogenous variables.

Two questions arise concerning the usefulness of the above model. First, are there reasonable or interesting economic hypotheses which imply belief formation functions which are linear in their parameters? Second, are there reasonable or interesting economic hypotheses which imply that the random elements enter in an additive manner?

---

(2) To establish identifiability we may assume we know the (population) joint distribution of $X_1, \ldots, X_n, Z_1, \ldots, Z_s$. In this distribution, equations (1) hold with $(\mathcal{E} X_j)$ substituted for X_j. (for, $\mathcal{E} u = 0$ and $\mathcal{E} \delta_{k\gamma} X_j Z_k Z_\gamma = \delta_{k\gamma} Z_k Z_\gamma \mathcal{E} X_j$). The question of identifiability is simply one of the usual rank criteria. If we assume that when the order criteria is satisfied the rank criteria is not likely to be unsatisfied, then we have a question of the number of known zeros which appear in these equations. The known zeros are more than enough to assure identifiability for many moments which appear in the system do not appear in the ith equation, and many of the Z's appear in only one or two moments but not in all. (Consider for example the variable "average earnings of the U.S. Steel Corporation"). There is one equation missing from the system. It states that the marginal rate of transformation, between mean and variance, is equal to the marginal rate of substitution of mean and variance. This depends on actual mean and variance which in turn depends on all the exogenous and endogenous variables. This missing equation, therefore, will not "spoil" the identification of the other equations.

We can not hope that all interesting or reasonable economic hypotheses imply belief formation functions which are linear in parameters. But some hypotheses do imply this. For example, classical security analysis gives us a number of variables (e.g., dividends per share, ratio of earnings to fixed interest payments, etc.) which are to be consulted in the evaluation of securities. We might hypothesize that the subjective probability beliefs of the investor are, roughly, "weighted averages" of the most important and relevant of such variables. This implies linear b.f.f's. It would be interesting to see how well this hypothesis works; first, because of interest in the hypothesis itself; and second, as a standard of comparison in judging the performance of more radical hypotheses.

The assumption of additive random variables is, on the other hand, extremely questionable. First we will present arguments which make it doubtful that an additive shock term enters our equations. Then we will consider hypotheses as to source of randomness and the implications of these hypotheses as to how random elements should enter our equations.

Equations (1) are known a-priori to be homogeneous. Somehow I would feel uncomfortable in assuming that such equations had "random constant terms."

Equations (1) were derived in Discussion Paper Economics 294. A different set of behavior equations, more useful for some purposes, less useful for others, was derived in Discussion Paper Economics 278. A third set of equations (which has not been presented) differs from the above and may prove useful when we do not know even the forms b.f.f's. Each of these sets of equations was derived from a common source, a fourth set of equations.

If we assumed that random elements entered additively in one of these four sets of equations, then we can not, with consistency, assume that it enters additively in the other three. It seems almost dishonest to presume that the particular set of equations which we use today is blessed with additive random elements, while the equations we derived yesterday or may use tomorrow are not so blessed.

Let us consider hypotheses as to the sources of randomness and their implications as to how random elements enter our equations.

1. One hypothesis is that equations (1) exactly determine a rational or optimal portfolio $X^o = (X_1^o, \ldots, X_n^o)$ but because the human being is an inaccurate or lazy mechanism the actual portfolio $X = (X_1, \ldots, X_n)$ deviates from the optimal by some random term $\epsilon = (\epsilon_1, \ldots, \epsilon_n)$. Our model then becomes (for suitably defined δ and Z)

$$(2.a) \qquad \sum_{j=1}^{n} \sum_{k=1}^{m} \delta_{ijk} X_j^o Z_k = 0 \qquad i = 3, \ldots, n$$

$$(2.b) \qquad X_j = X_j^o + \epsilon_j \qquad\qquad (2.c) \quad \Sigma X_j = 1$$

where X_j Z_k are the observed variables; ϵ_j presumably has the properties $\mathcal{E} \epsilon_j = 0$, $\mathcal{E} X_i^o \epsilon_j = 0$, $\mathcal{E} Z_i \epsilon_j = 0$. The statistical problem, if we make this hypothesis, is to estimate δ_{ijk} and the distribution of the ϵ's.

2. We may hypothesize that the probability beliefs of equations (1) are random variables. In particular:

A. We may believe that the weight attached to some factor Z may vary from case to case and time to time. This variation, we may hypothesize, can be accounted for by assuming that the α's and the β's are random variables presumably independent of the Z's. Our model, then, is as in equations (1) above, but with the δ as random variables

$$\delta = \alpha\beta = (\alpha^o + \epsilon)(\beta^o + \eta) = \alpha^o \beta^o + \beta^o \epsilon + \alpha^o \eta + \epsilon\eta = \delta^o + \xi$$

$$(\alpha^o, \beta^o, \delta^o \text{ constants})$$

The problem is to estimate δ^o and the distribution of the ξ's.

B. We may believe that fluctuations in subjective probability beliefs are due to variations in "Z's" which were not taken into account in our analysis. Thus the "truth" may be

$$(3) \qquad \mu_i = \alpha_o + \alpha_1 Z_1 + \cdots + \alpha_S Z_S + \alpha_{S+1} Z_{S+1} + \cdots + \alpha_t Z_t$$

while our analysis takes into account only the first s Z's and ascribes variation in the rest to a random variable η

(4) $$\mu_1 = \alpha_0 + \alpha_1 z_1 + \cdots + \alpha_s z_s + \eta$$

The difference between 2A and 2B is that in 2B the b.f.f's have "a random constant term," while in 2A we have "random coefficients." The model in case 2B has equations of the form

$$\Sigma \delta_{ijk} x_j z_k + x_j \eta = 0$$

The problem is to estimate the δ's and the distribution of the η's.

3. The hypotheses in 1, 2A and 2B all seem reasonable. All three sources of randomness may be hypothesized thus giving the equations

$$\Sigma(\delta^o + \xi) x^o z + x_j^o \eta_j = 0$$

where we observe Z and $X_i = X_i^o + \epsilon_i$. There are, of course, other hypotheses as to sources of randomness. We might hypothesize, for example, that our observations of, or knowledge about, some Z's differ from that of the investor. Or we may find reasons for doubting the independence of the Z's and the random variables. But the attention of statisticians is a scarce resource; the problems of 1, 2 and 3 seem to be of prime importance and are not likely to find a quick and simple solution.

We must now consider our results when utility depends on other moments besides mean and variance. In Discussion Paper Economics 294, p. 4 we describe the modifications in our behavior equations as more moments are added to the utility function. For suitably defined δ and Z we have, as the ith equations,

$$\sum_{j \ldots, n} \delta_{i,\ldots,m,n} X_{j,\ldots,} X_m \; Z_{i,j,\ldots,m,n} = 0 \,.$$

The same arguments concerning identifiability and estimation in the "shock model" case, the same arguments against the shock model and for models derived from the hypotheses in 1, 2 and 3 above, can be applied in the general case as were applied in the mean-variance case.

We must not expect that nonadditiveness of random variables will be peculiar to investment company behavior equations. In almost every business, management attempts to allocate funds among alternatives with uncertain outcomes in an effort to obtain desired ends concerning the returns from the business as a whole. Complications, not found in the study of open-end investment company, may arise because of nonlinear production functions, fixed obligations, differences of interest between management and various classes of security holders — to name a few which come to mind. The stochastic models which "fit" the investment company, will typically not "fit" other institutions. It is hard to believe, however, that when we add these institutional complications our nonadditive random variables will somehow "disintangle" themselves and the appropriate models will be shock models. We must expect, I fear, that shock models have little or no place in the econometrics of business behavior.

Part B

In Part A of this paper we presented some statistical problems. In this part of the paper we present a model ready to be "tried out" empirically. This model makes the assumptions of Cowles Commission Discussion Paper Economics 278, Part 2 (or, equivalently, Discussion Paper Economics 294). This implies that (1) the investigator must specify which moments are assumed to enter the investor's utility function, and (2) he must specify at least the forms of the b.f.f.'s.

Partly because of the difficulties raised in Part A of this paper, partly because of the nonlinear nature of the hypothetical b.f.f's, the investigator will specify the b.f.f's completely. That is, he will leave no unknown parameters to be estimated. The proposal is to derive the "efficient surface" using these b.f.f's, and then compare this surface with the actual portfolios of various investment companies.

We will assume that utility is a function only of expected return and variance. Reasons for this decision are given in the footnote.(3)

(3) In deciding which hypothesis to assume we must consider the costs of and returns from each decision. Let us begin, then, by considering the costs involved in adding the third (or higher) moments to the utility function. The mean-variance hypothesis will involve inverting and operating with linear equations. If we add the third moment μ_3 then we have third degree equations. The hypothesis that the first n moments enter the utility function gives rise to equations of degree $\frac{(n-1)n}{2}$. Besides the increased complexity of the equations, there are the additional moments (concerning the flow of returns from individual securities) which must be computed. Although our equations will take account of only a few securities, marginal computation costs are large and must be taken into account.

As more moments are added to the utility function the dimensionality of the efficient surface increases; the "strength" of the hypothesis decreases.

We will see later that we have available to us, when we make the mean-variance hypothesis, a method of presentation which is conducive to the intuitive appreciation and critical evaluation of our final results. This method is not as useful when more than two moments are assumed to enter the utility function. In sum, there is marginal cost involved in adding more moments to the utility function. There are costs in terms of (1) greatly increased computation time and expense, (2) loss of information yielded by the hypothesis if it is true (i.e., loss of strength), and (3) loss of intuitive grasp of final results.

The Third moment (μ_3) is commonly associated with the "propensity to gamble." (Note e.g., that if an individual has a Bernoulli utility function:

$$U = \alpha + \beta Y + \gamma Y^2, \quad \beta > 0, \quad \gamma < 0,$$

then he will never take a fair bet; if

$$U = \alpha + \beta Y + \gamma Y^2 + \delta Y^3, \quad \delta > 0,$$

then there are some fair bets which he will take). This "propensity to gamble" seems important in the explanation of the behavior of some investment companies (closed-end companies with "leverage"; speculative open-end companies). In other companies (those whose aim is income and stability of income) this propensity to gamble may be small. If we consider only those companies which express, or are reputed to have, such goals, then the mean-variance hypothesis may not be "far off." More explicitly, if $U = U(E, V, \mu_3)$ then we have

$$\frac{\partial U}{\partial E}\frac{\partial E}{\partial X_i} + \frac{\partial U}{\partial V}\frac{\partial V}{\partial X_i} + \frac{\partial U}{\partial \mu_3}\frac{\partial \mu_3}{\partial X_3} + \lambda = 0$$

$$i = 1, \ldots, n \qquad \sum X_i = 1$$

If $\frac{\partial U}{\partial \mu_3} \equiv 0$ then we have the mean-variance case. If $\frac{\partial U}{\partial \mu_3}$ is relatively small, then the actual allocation will be close to the mean-variance efficient surface. The neglect of μ_3 will be one of the many factors which make the observed portfolios deviate from the hypothetical efficient surface. The fourth and higher moments seem to affect action because of motives already taken into account by E, V and μ_3. We may hypothesize that they are "almost" superfluous.

In sum, if we restrict our considerations to certain kinds of investment companies, the marginal return due to adding μ_3 (or other moments) to the utility function may not be great. Since it appears that $\Delta C > \Delta R$, we do not take the marginal action.

The equations to be fit will take into account only the securities of one industry (probably the common stock issues of the steel industry). Of the stocks of this industry, only those will be considered which are held by some investment company to be used in the testing of the hypothesis. The justification for the latter exclusion flows from the fact that if the point $(X_1, \ldots, X_n, 0, 0, \ldots, 0)$ is efficient, then so is the point $(X_1, \ldots, X_n)$. The justification for restricting consideration to the common stocks of one industry is the discussion of segregate equations in Discussion Paper Economics 294, p. 5. There we saw that if we assume $\rho_{ik} = \rho_{jk}$ for all i, j, k where i, j, refer to securities within the industry and k to securities outside the industry, then equations could be found concerning the allocation of funds among securities of a single industry. We also considered a model showing some conditions under which $\rho_{ik} = \rho_{jk}$.(4) The steel industry seems to meet these conditions as well as most industries.

The segregate efficient surface is a function of $K = \sum_{j=k+1}^{n} \rho_{ij} \sqrt{\sigma_{jj}} X_j$ where securities $1, \ldots, k$ are within the industry; $k+1, \ldots, n$, outside the industry; i is any number $\leq k$. K could be computed for each investment company using the b.f.f's described below. Considerations of computation time rule this out. The K of various investment companies can be estimated from a random sample of the ρ_{ij}, $\sqrt{\sigma_{jj}}$, X_j. We can derive the efficient surface for one or more typical values of K.

Suppose we have calculated our hypothetical μ_i, σ_{ij} and have estimated the hypothetical K. Suppose that from these numbers we have calculated the efficient set and the attainable combinations of E and V. How do we compare our theoretical results

---

(footnote (3) continued) Here again we have sought subject matter which give rise to simple behavior equations. The justification for such action lies in the nature and goals of this project.

(4) We will exclude from our analysis any investment companies which also hold the preferred stocks or bonds of the steel industry. Such holdings are not common. The exclusion of such investment companies is required by our segregative assumptions.

with actual portfolios? One possibility is to somehow compare the actual allo- cations $(X_1,\ldots,X_n)$ with the theoretically efficient ones. A drawback to such an ann approach is that the dimensionality of the allocation space is too great to permit an intuitive grasp and evaluation of the results. Another approach is to compare the actual combinations (E,V) with the theoretically efficient set, assuming our μ_i, σ_{ij}, K are correct. The final result may look like Figure 1. If the μ, σ, K were correct and our hypothesis held exactly, then all observed points would lie on the efficient segment. If the actual points are "near" the efficient segment then the hypothesis is "good" and may be "useful."

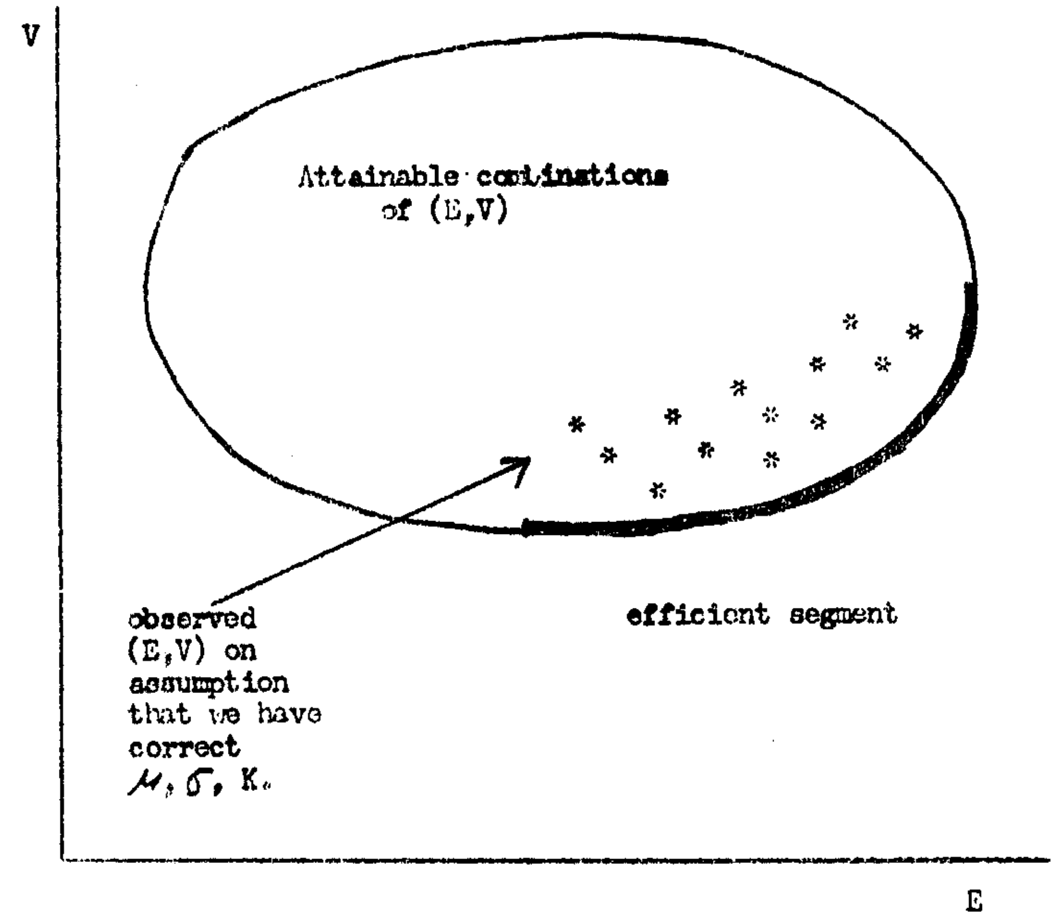

Figure 1

Given μ_i, σ_{ij} and K, we have the problem of computing the efficient surface and the attainable combinations of E and V. This is a problem of finding the extreme values of $V = \sum_{i,j=1}^{k} \sigma_{ij} X_i X_j + K \sum_{i=1}^{k} \sqrt{\sigma_{ii}} X_i$ given $E = \sum \mu_i X_i$, $\sum X_i = 1$, $X_i \geq 0$. For given E we can solve this problem by gradient methods. For a given set of securities $g_1, \ldots, g_\alpha$, we can solve the problem for all extreme points such that $X_{g_1} > 0, \ldots, X_{g_\alpha} > 0$, $X_i = 0$ for all $i \neq g_1, \ldots, g_\alpha$. (See Discussion Paper Economics 278, Part II). In principle we could find the attainable points (E,V) by computing the extreme points for all combinations of securities $g_1, \ldots, g_\alpha$. But this is impractical. We will have to rely heavily on extrapolation unless superior computational procedures become available.

The Belief Formation Functions

We have discussed what we can do when we have the hypothetical μ_i and σ_{ij}. Now we will present our b.f.f's. We will not give extended and detailed justifications for the various economic assumptions made. Such justification must be made at some time; but it seems inappropriate in this paper with its emphasis on statistical and computational problems. We will give, however, a general discussion of (a) what we can expect of b.f.f's in general; and (b) a fundamental weakness of our whole analysis. We then present our b.f.f's with only as much justification as is needed to motivate our discussion.

(a) We cannot expect our b.f.f's to be either "true" or "rational." They are neither the procedures which investment companies actually use in forming subjective probability beliefs, nor the procedures which they ought to use in forming such beliefs. Both the "true" and "rational" b.f.f's (whatever they be) must surely take into account more variables than can be considered in most econometric studies. Our goal is to find b.f.f's which give, in some sense, plausible results using a few, commonly available variables.

(b) The basic assumptions of Discussion Paper Economics 278, Discussion Paper Economics 294 and this paper are static assumptions. We have assumed that the investor has probability beliefs about "the flow" of returns from each individual security, and that he maximizes utility which depends on the probability distribution of "the flow" of returns from the portfolio as a whole.

The chief reasons for making these static assumptions were: (a) it was presumed that dynamic assumptions would lead to equations describing optimal strategies rather than optimal allocations. Behavior equations describing optimal strategies seem to be, at present, completely unmanageable computationally. (For example, see Discussion Paper Economics 295, p. 13. These are not investment company behavior equations. But similar techniques can be used and similar results obtained.); (b) investment behavior, as distinguished from speculative behavior, is traditionally described in terms of the allocation of funds among securities to be held for their yield. Investment is an allocation problem; speculation, a problem of strategy. One might argue that, for nonspeculative investment companies, static assumptions may not be far wrong; (c) static approximations to dynamic realities have sometimes proven extremely useful. For example, probability is essentially a static concept customarily applied to properties of sequences of events.

For these reasons we made our static assumptions. These assumptions give rise to the problem of describing a dynamic reality by a static approximation. In particular we must anticipate that expected returns and covariances of returns will be a function of time ($\mu_i(t)$, $\sigma_{ij}(t)$). We must substitute single numbers μ_i, σ_{ij} for these functions. It seems natural to discount $\mu_i(t)$ by some discount rate r. This we will do; but what value of r shall we use? It may be desirable to carry out more than one analysis using different plausible values of r. It does not seem as natural to discount $\sigma_{ij}(t)$. Several alternate possibilities were considered, but nothing conspicuously better could be found. It was therefore decided to "discount" the $\sigma_{ij}(t)$ by a factor s. It may be desirable to carry out more than one analysis using different values of s. (An important difference between K on the one hand, and r and s on the other is that K is estimated; r and s are guessed at.)

There is a more fundamental weakness in the use of r and s than the fact that their values must be guessed at and the discounting of covariances is unfamiliar. Subjective discount factors are presumably matters of preference rather than of opinion. $(1+r)$ is presumably the rate at which the investor is just willing to trade future for present $\mu_i(t)$. μ_i and σ_{ij} (concerning "the flow" of returns) are, on the other hand, supposed to be matters of opinion rather than of preference.

These difficulties are not due to some superficial decision of ours. They are due to our underlying static assumptions. These difficulties, therefore, suggest that our theory should be completely reconstructed. I believe they also suggest the way the theory (and, from the theory, the behavior equations) should be reconstructed. But this is another story.

The point is that we are in what some people would consider a very uncomfortable situation. Even before our hypothesis is tested empirically, basic theoretical weaknesses call for its complete reconstruction. I suspect that after careful consideration, our reconstructed theory will require re-reconstruction; the re-reconstructed theory, further reconstruction and this will continue as long as progress is to be made.

At each stage, with each new approach, we must subject our best hypotheses to empirical test. We must "try them out" on actual data. This is necessary for two reasons: (1) Although it may be subject to theoretical difficulties, our theory may work well enough to be useful. (2) The empirical tests may suggest new difficulties to be taken into account. It may be that with minor modifications we can produce a theory that "works." Or the evidence may require and contribute to a more fundamental reconstruction. It is in this spirit that we consider the theory at hand.

* * *

We now present our hypothetical b.f.f's.

Our final μ_i and σ_{ij} will combine two sorts of information: (a) what expected dividend from stock i, and covariance of dividend between i and j, should we anticipate at time t if we were sure that corporations i and j will not have failed by time t? (b) what is the probability that i will have failed, what is the probability that both i and j will have failed by time t?

Concerning (a) we take the following steps:

(1) Draw a linear (least squares) trend through dividends. Our series must begin and end in the same phase of the business cycle. Let $t = 0$ be today. The fitted trend is $d^o_{i,t} = a_i + b_i t$.

(2) We assume that the standard deviation of dividends is proportional to its expected value; i.e., $d_{it} = \mathcal{E} d_{it} + (\mathcal{E} d_{it}) u_i$ where u_i has constant variance $V u_i$. Estimate $V u_i$ by

$$s_i^2 = \frac{1}{n-2} \sum \left(\frac{d_{it} - d^o_{it}}{d^o_{it}} \right)^2 .$$

Estimate the covariance of u_i and u_j by

$$c_{ij} = \frac{1}{n-2} \sum \frac{(d_{it} - d^o_{it})}{d^o_{it}} \frac{(d_{jt} - d^o_{jt})}{d^o_{jt}} .$$

We have assumed that the correlation between d_i and d_j is constant through time.

$$\rho_{d_{it}\, d_{jt}} = \frac{c_{d_{it}\, d_{jt}}}{s_{it}\, s_{jt}} = \frac{c_{ij}\, d^o_{it}\, d^o_{jt}}{s_i\, d^o_{it} \cdot s_j\, d^o_{jt}} = \frac{c_{ij}}{s_i\, s_j} .$$

d^o_{it}, $s_{it} = s_i d^o_{it}$, $C_{ijt} = C_{ij}\ d^o_{it}\ d^o_{jt}$, are the anticipated moments at time t, assuming that corporations i and j will survive till then.

Concerning (b) take the following steps:

(3) Draw least squares trend line through "gross earnings" (earnings before depreciation, taxes or bond interest are subtracted). Let $t = 0$ be today. The fitted trend in $g^o_{it} = \alpha_i + \beta_i\ t$.

(4) Make the same assumptions, calculate the same estimates of variances and covariances of "gross earnings," as were prescribed in (2) for dividends. Those means, variances and covariances of gross earnings at time t, assuming that corporations i and j are still in existence then, will be labled

$$s^g_{it} = s^g_i\ g^o_{it} \qquad C^g_{ijt} = C^g_{ij}\ g^o_{it}\ g^o_{jt}\ .$$

(5) Let B_i be the fixed interest payments of corporation i. P_i will be called the probability of failure of corporation i. It will be estimated by the probability that a random variable distributed normally with zero mean and unit variance will fall below $\dfrac{B_i - \alpha_i}{s^g_i\ \alpha_i}$. P_{ij} will be called the probability of both i and j failing. It will be estimated by the probability that two normally distributed g_i, g_j variables with means α_i, α_j and covariances $(s^g_i\ \alpha_i)^2$, $(s^g_j\ \alpha_j)^2$, $C^g_{ij}\ \alpha_i\ \alpha_j$ will fall below B_i and B_j respectively. Let $q_i = 1 - P_i$; $q_{ij} = 1 - P_{ij}$.

Now we must combine the results of steps 1 - 5.

(6) The probability that corporation i will not fail by time t is $(1 - P_i)^t = q^t_i$. The expected dividend if the corporation survives is $d^o_{it} = a_i + b_i\ t$. The expected dividend, taking into account the probability of failure, is $D_{it} = q^t_i\ (a_i + b_i\ t)$. The anticipated covariance, taking into account the probability of failure, is

$$\begin{aligned} M_{ij,t} &= (q_{ij})^t\ \mathcal{E}(d_i\ d_j) - (q_i \cdot q_j)^t\ \ d^o_{it}\ d^o_{jt} \\ &= (q_{ij})^t\ (C_{ij} + 1)\ d^o_{it}\ d^o_{jt} - (q_i \cdot q_j)^t\ d^o_{it}\ d^o_{jt}\ . \end{aligned}$$

(For $i = j$ we have the variance of i).

(7) The discounted expected values and variances are, respectively,

$$D_i = \sum_{t=1}^{\infty} \frac{D_{it}}{e^{rt}} \qquad\qquad M_{ij} = \sum_{t=1}^{\infty} \frac{M_{ijt}}{e^{rt}}$$

We can not consider the discounted values D_i, M_{ij}, as the expected values and co-variances of "the flows" of returns. Suppose for example that D_{it} is a constant (a) through time. Then

$$D_i = \sum_{i=1}^{\infty} \frac{a}{e^{rt}} = a\left(\frac{1}{e^r - 1}\right)$$

[C.f. Step 9 below, with $b_i = 0$, $q_i = 1$]. But surely the expected value, as distinguished from the discounted value, of the flow is a. We therefore define the "expected value of the flow" as $\mu_i = (e^r - 1)\, D_i$. Similarly we define the co-variances of the flow as

$$\sigma_{ij} = (e^r - 1)\, M_{ij}.$$

(8) D_i and M_{ij} can be evaluated with the help of the following infinite sums.

i) $$\sum_{t=1}^{\infty} \frac{q^t}{e^{rt}} = \sum_{t=1}^{\infty} \left(\frac{q}{e^r}\right)^t = \frac{q}{e^r}\left\{\sum_{t=0}^{\infty} \left(\frac{q^t}{e^r}\right)\right\} = \frac{q}{e^r}\ \frac{1}{1 - \frac{q}{e^r}} = \frac{q}{e^r - q}$$

ii) $$\sum_{t=1}^{\infty} \frac{q^t \cdot t}{e^{rt}} = -\sum_{t=1}^{\infty} \frac{\partial q^t e^{-rt}}{\partial r} = -\frac{\partial \frac{q}{e^r - q}}{\partial r} = \frac{q\, e^r}{(e^r - q)^2}$$

iii) $$\sum_{t=1}^{\infty} \frac{(q)^t\, t^2}{e^{st}} = \sum_{t=1}^{\infty} \frac{\partial^2 (q)^t e^{-st}}{\partial s^2} = q\ \frac{\partial \frac{-e^s}{(e^s - q)^2}}{\partial s} = q\left(\frac{(e^{2s} + q)}{(e^s - q)^3}\right)$$

(9) In sum, our belief formation functions are as follows:

$$\mu_i = (e^r - 1)\, D_i = (e^r - 1)\left\{ a_i\left(\frac{q_i}{e^r - q}\right) + b_i\left(\frac{q_i\, e^r}{(e^r - q_i)^2}\right)\right\}$$

$$\sigma_{ij} = (e^s - 1)\, M_{ij} = (e^s - 1) \Bigg\{ a_i\, a_j \left[C_{ij}\ \frac{q_{ij}}{e^s - q_{ij}} + \frac{q_i\, q_j}{e^s - q_i\, q_j} \right]$$

$$+ (a_i\, b_j + a_j\, b_i) \left[C_{ij}\ \frac{q_{ij}\, e^s}{(e^s - q_{ij})^2} - \frac{q_i\, q_j\, e^s}{(e^s - q_i\, q_j)} \right]$$

$$+ b_i\, b_j \left[C_{ij}\, q_{ij} \left(\frac{-(e^{2s} + q_{ij})}{(e^s - q)^3} \right) + q_i\, q_j \left(\frac{-(e^{2s} + q_i\, q_j)}{(e^s - q_i\, q_j)^3} \right) \right] \Bigg\}$$

Where a_i, b_i, a_j, b_j, C_{ij}, q_i, q_{ij} are as defined in 1-5.