

INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.
2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.
4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.
5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms

300 North Zeeb Road
Ann Arbor, Michigan 48106

73-21,916

MILLER, Leonard Allen, 1943-

DIMENSIONS OF PORTFOLIO COMPOSITION: AN APPLICATION OF MULTIDIMENSIONAL SCALING.

The City University of New York, Ph.D., 1973
Business Administration

University Microfilms, A XEROX Company, Ann Arbor, Michigan

DIMENSIONS OF
PORTFOLIO COMPOSITION

An Application of
Multidimensional Scaling

by

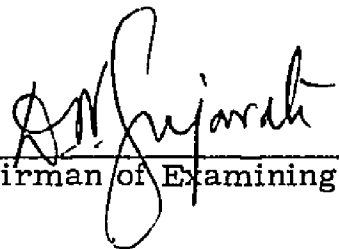
Leonard A. Miller

A dissertation submitted to the Graduate
Faculty in Business Administration in
partial fulfillment of the requirements
for the degree of Doctor of Philosophy,
The City University of New York.

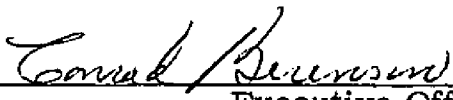
1973

This manuscript has been read and accepted for the Graduate Faculty in Business Administration in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

5/15/1973
date


Chairman of Examining Committee

5/15/73
date


Executive Officer

Supervisory Committee:

Professors: Damodar Gujarati
Ronald Gatty
Lloyd Rosenberg
Martin Zweig

ACKNOWLEDGMENT

The list of people and organizations who directly or indirectly made this dissertation possible is a long one, reflecting the lengthy educational process that necessarily preceded it. Of those most directly connected, I wish to thank Dr. Damodar Gujarati and the other members of my Supervisory Committee for their guidance. Also, Dr. Vithala Rao of Cornell University is thanked for his constructive review.

Dr. Peter Gutmann, Chairman of the Economics Department at Bernard M. Baruch College, and Dr. David Valinsky, Chairman of the Statistics Department, have my deepest appreciation. The concern they showed for the success of individual students became, in turn, a motivation to succeed.

I am indebted to the staff of the Baruch College Computer Center and the staff of the City University Graduate Division Computer Center. They provided professional assistance beyond normal expectations in what turned out to be a truly massive data processing task.

The Arthur Lipper Corporation provided invaluable data. Their generosity and expeditiousness in responding to repeated requests is gratefully acknowledged.

My gratitude is extended to the Bell Telephone Laboratories and the University of Pennsylvania for the assistance they provided. Dr. J. Douglas Carroll of Bell Laboratories deserves my special thanks. He inspired this dissertation and gave helpful instruction.

Finally, my greatest debt is to my wife, Teresa. She deserves an important share of the recognition for the dissertation--for her cheerful patience and encouragement through long years of schooling as well as her expert secretarial assistance. This dissertation is dedicated to her.

— Leonard Miller

TABLE OF CONTENTS

<u>Chapter</u>		<u>Page Number</u>
1	INTRODUCTION: A CONCEPTUAL AND HISTORICAL FOUNDATION	1
2	AN EXPERIMENTAL DESIGN: DATA PREPARATION	48
3	THE EXPERIMENT	62
4	SCENARIO DEPENDENCE OF INVESTMENT PREFERENCES	96
5	EXPERIMENT DESIGN VARIATIONS	110
6	SUMMARY	135
	APPENDIX	139
	BIBLIOGRAPHY	158

LIST OF TABLES

<u>Table</u>		<u>Page Number</u>
1	Consensus Mutual Fund Classification System By Investment Objective Categories	51
2	Population of Mutual Funds as of December 31, 1970	52
3	M-D-SCAL Experiment Stress Versus Dimensionality	66
4	Stress Quality	68
5	Degree of Correspondence Between Natural and Predicted Clustering for Two Dimensional M-D-SCAL Solution	77
6	Degree of Correspondence Between Natural and Predicted Clustering for Three Dimensional M-D-SCAL Solution	91
7	INDSCAL Configuration Matching	104
8	Stress Versus Dimensionality for Scaling Solutions Based on Alternative Proximity Measures	112
9	Stress of Scaling Solutions Within City-Block Metric Space for Alternative Proximity Measures	121
A1	Experiment Sample of 46 Mutual Fund Subjects	139
A2	Standard Investment Groups	141
A3	Mutual Fund Portfolio Holdings Percent of Total Assets	143
A4	Holdings Overlap Similarities Between Mutual Fund Portfolios	145
A5	Alternative M-D-SCAL Solutions--March 31, 1971	146

LIST OF TABLES (Continued)

<u>Table</u>		<u>Page Number</u>
A6(a)	Hierarchy of Natural Clusters in the Two Dimensional M-D-SCAL Mutual Fund Solution Space	147
A6(b)	Hierarchy of Natural Clusters in the Three Dimensional M-D-SCAL Mutual Fund Solution Space	147
A7	Mutual Fund Yield and Growth Property Vectors	148
A8	Alternative M-D-SCAL Solutions--September 30, 1971	149
A9	Alternative M-D-SCAL Solutions--June 30, 1971	150
A10	Alternative M-D-SCAL Solutions--December 31, 1971	151
A11	Property Fit of Investment Group Vectors	152
A12(a)	MDPREF Solution--Stimuli Points Investment Group Coordinates	155
A12(b)	MDPREF Solution--Subject Vectors Mutual Fund Direction Cosines	157

LIST OF ILLUSTRATIONS

<u>Figure</u>		<u>Page Number</u>
1	Farrar's Mean-Variance Feasible Region Showing Actual, Optimal and Randomized Portfolios	22
2	M-D-SCAL Mutual Fund Solution Space Based Upon Holdings Overlap Similarities as of March 31, 1971	65
3	M-D-SCAL Solution Stress for Alternative Dimensionalities	67
4	M-D-SCAL Solution Space--March 31, 1971 Showing Clustering of Growth and Income Oriented Mutual Funds	70
5	M-D-SCAL Solution Space--March 31, 1971 Showing Mutual Funds Clustered in Accordance With Their Investment Objective Classifications	72
6	Natural Clusters at the Two Group Level in the Two Dimensional M-D-SCAL Mutual Fund Solution Space	74
7	Natural Clusters at the Four Group Level in the Two Dimensional M-D-SCAL Mutual Fund Solution Space	75
8	M-D-SCAL Solution Space Showing Centroids of Mutual Fund Clusters	79 ^a
9	Mutual Fund Yield and Growth Property Vectors Fitted into the M-D-SCAL Solution Space	83
10	M-D-SCAL Solution Space Rotated 29 Degrees Clockwise	85
11	M-D-SCAL Mutual Fund Solution Space--June 30, 1971, With Enclosed Investment Objective Groups	98
12	M-D-SCAL Mutual Fund Solution Space--September 30, 1971, With Enclosed Investment Objective Groups	99

LIST OF ILLUSTRATIONS (Continued)

<u>Figure</u>		<u>Page Number</u>
13	M-D-SCAL Mutual Fund Solution Space-- December 31, 1971, With Enclosed Investment Objective Groups	100
14	M-D-SCAL Mutual Fund Solution Space Based Upon Euclidean Dissimilarities	114
15	Investment Group Vectors Fitted Within M-D-SCAL Mutual Fund Solution Space	126
16	MDPREF Joint Space Solution Showing Investment Group Points and Mutual Fund Vectors	129

Chapter 1

INTRODUCTION: A CONCEPTUAL AND HISTORICAL FOUNDATION

The population of investors consists of many varieties. Investors may be classified according to the type of assets they acquire--industry and government securities, real estate, commodities, art work and so forth. A taxonomic distinction also may be made between those investors who select among individual assets and build portfolios and, at a different level, those who invest in portfolios already constructed.

A central issue in the modern theory of finance has been the question: How to construct an optimal portfolio from the universe of available individual investments. This question is the crucial concern for those investors who are portfolio architects, especially professional institutional money managers.

A different and relatively neglected question is raised by those who are investors in portfolios, particularly shareowners of investment companies such as mutual funds. That is, how to identify the portfolio from the set of already constructed portfolios which best meets a given set of investment objectives? It is this question to which the present study is related.

The problem of selecting which is the best mutual fund portfolio for a given investor is not as simple as it might seem. The notion of

"best" implies some criterion or yardstick of investment objectives. Certainly if we are willing to wait for the outcome, the issue of "best" is easy but moot; it is the portfolio giving the largest terminal dollar reward. However, when the question is posed in a prospective sense, two questions of fundamental theoretical and practical importance arise:

- 1) What are the investment objective variables by which investors discriminate and select among alternative portfolios? In other words, what are the portfolio attributes by which investors perceive important differences between portfolios?
- 2) If we can establish what are the basic factors of portfolio discrimination, how can we scale existing portfolios in terms of these variables? That is, how can we measure portfolios along each dimension of preference so as to provide investors information they desire to reach decisions.

Conventional approaches to answering these questions are found to be wanting. Usually the assumption is made that the psychological dimensions of portfolio evaluation are those associated with modern utility theory, namely, expected return and some index of risk. An individual portfolio is scaled with respect to these variables by measurement of the portfolio's historical behavior. This approach is unsatisfactory on two counts:

- 1) The theoretical dimensions of evaluation remain as assumptions about what actually are the psychological variables that investors invoke to discriminate between portfolios. This represents making a presumptive equivalence between "what is" and "what should be."

- 2) The use of historical data to measure investment objectives is not meaningful.

With regard to the last point, expected return and risk are ex ante concepts. However, because of the difficulty of obtaining useful opinions as to expectations and to meet the demands for "objectivity," ex post measurements usually are taken as surrogates for ex ante judgments. And the measurements are often made in terms of a share of the mutual fund itself rather than the individual investments which comprise the fund's portfolio. Thus, a probability distribution of the historical price movements of a fund share is established and the mean and variance (or standard deviation) of the distribution are taken to represent the fund's expected return and risk, respectively. But clearly, focusing on the past behavior of the fund in this manner tends to make it impossible to detect shifts in the fund's investment objective over time engendered, say, in response to swings in the economic environment. The current emphasis of investment objectives for a given mutual fund, as reflected in the present structure of its portfolio, is not conveyed.

To avoid this problem, ex post risk and return values may be built up as a composite of the historic values for the component assets which currently comprise the fund's portfolio. Nonetheless, well-known propositions of efficient capital markets and random walk price movements, supported by the evidence of experience, makes the use of

historical measurements of behavior in any form an unsatisfactory way, indeed, of characterizing the current orientation of investment objectives for a given mutual fund.

Instead of relying on explicit measurements of past behavior, portfolios can be compared in a prospective sense in terms of the subjectively stated objectives of the mutual funds which hold the portfolios. By law each mutual fund must proclaim its investment objectives. Following are representative statements by actual mutual funds:

- "Broad Street Investing's primary investment objective is to produce favorable current income and long-term growth of income, with due regard to preservation of capital value. Long-term growth of capital value also is an important goal." [1]
- "Whitehall Fund seeks to provide investors with a prudent program of 'balanced investing for gain.' A portion of the Fund's assets is invested in bonds and preferred stocks to provide basic stability of capital value and to buttress income. Assets are invested in common stocks to provide possibilities of growth of capital value and income." [2]

Needless to say, statements such as these are subject to broad interpretation and do not provide the individual investor a firm basis for close comparison. To assist the prospective mutual fund investor, various organizations provide the service of classifying funds. Generally, the classification schemes have their origins in the conventional wisdom that growth (capital appreciation) and income (dividends and interest) are the important dimensions of discrimination with stability

(risk) as an ancillary consideration. Using these descriptors, a hierarchy of categories can be defined, e.g., Growth, Growth-Income, Growth-Income-Stability, Income-Growth, Income. Assigning a given fund to any such category is based upon consideration of the fund's self-stated objectives and policies, its past behavior as well as an appraisal of the characteristics of the fund's current holdings.

Although perhaps providing more mature information as to a fund's real investment objectives than does a simple measurement of the fund's past behavior, reliance on this professional-appraisal approach still is found to be wanting. That the problem of subjective interpretation is not cured is evidenced by the fact that despite the broadness of the categories, it is not infrequent for different classifying services to assign a given fund to different categories. Even more important, a classification scheme based upon categories alone represents a nominal scale of investment objectives, a very weak form of information.¹ Differences between funds within a category are not conveyed.

In light of the foregoing, there exists need for an improved method of rating mutual fund portfolios in terms of the basic factors by which investors perceive differences between investments. Ideally, the method would be "current" in the sense of not relying on data concerning

1. More properly, conventional mutual fund classification systems represent a hierarchical form of nominal scale. Chapters 2 and 3 discuss in more detail these mutual fund classifying services and their classification practices.

the past behavior of the fund itself or its component assets. It would be "objective" in the sense of not involving the gathering and interpretation of subjective anticipations (opinions) of future behavior. It would be informative by presenting metric scaled differences between portfolios along the dimensions of perception instead of just giving categorical differences. It would be "sensitive" by being able to detect shifts in a fund's investment objectives as reflected in periodic changes in the mix of its portfolio holdings. Finally, it would be "realistic" in the sense that it would not require the prior assumption that certain measurable portfolio attributes correspond to the basic perceptual dimensions. Instead, the method itself will evince the dimensions which account for differences between portfolios and simultaneously, will scale each mutual fund portfolio with respect to these found dimensions.

This report reflects a presentation of a new method which is designed to satisfy these requirements. The method is rooted in the notion that there exists sufficient informational content in a breakdown of the holdings of a large sample of mutual funds to: 1) evince general factors which can account for differences in holdings, and 2) scale metrically the portfolios with respect to these uncovered basic factors. The method is made possible by the development of a new statistical technology, nonmetric multidimensional scaling.

The aim of this study may be regarded as a quest for a positive model of investment evaluation--a description of how investors do

perceive differences between portfolios--rather than a normative model for "rational" selection--how they ought to perceive differences. The pursuit of a description of real behavior does not imply the belief that normative theorizing is inappropriate or less important. This study is motivated by the belief that the plausibility of any normative model of human behavior is how well it reconciles with what is known to be the real human factors of choice and preference. Only with respect to these true factors of human preference can the superiority of a normative model be determined relative to actual behavior. Knowledge of the real factors of preference must come first.

Dimensions of Investment Evaluation Theory and Practice

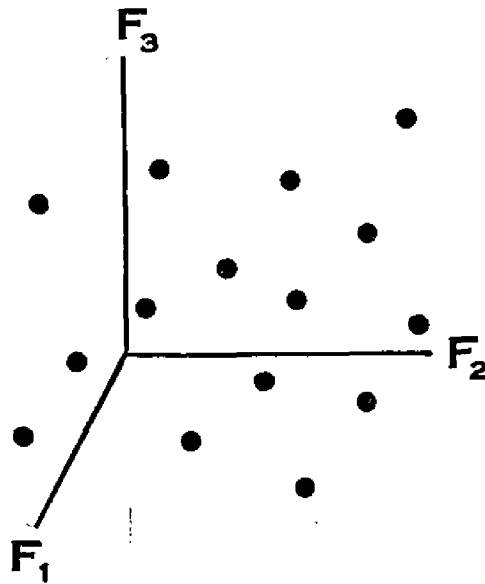
The familiar concept of a utility function may be regarded as a model of the psychological process of evaluation. In this sense, "value" becomes a hypothetical construct as the compound of various perceived effects. In line with empirically and axiomatically [3] developed theories of human perception, a strong case can be made that investors' utility functions are unique up to a linear transformation. That is, the utility of an asset is an additive function of its various investment attributes; namely -

$$U_{sj} = a_{1s} f_{1j} + a_{2s} f_{2j} + \dots + a_{ns} f_{nj}$$

where U_{sj} stands for subject S's utility for stimuli J, say portfolio J. f_{1j} through $f_{n-1 j}$ are the characteristic scale values of item J on those

dimensions of perception, $F_1, \dots, F_{n-1}$, which are common to all investors. f_{nj} is the scale value on a perceptual dimension, F_n , which summarizes one or more perceptual dimensions that are unique to subject S. The $a_1, \dots, a_n$ coefficients are the subject's idiosyncratic weightings of the various perceptual dimensions.

This model of evaluation implies for each investor a mental multi-dimensional "perceptual map" of the stimuli where each portfolio is perceived as a point in a space spanned by coordinates that correspond to the arguments of the investor's utility function. Graphically, for the three-factor case:



The idea is that each portfolio evokes responses on the different preference attribute scales. The projections of a portfolio point onto

the axes of perception represent these responses. Depending upon the nature of the preference variables, the responses may be metric scaled or perceived only in rank order terms.¹

Conventional normative models of investment selection posit the basic dimensions of perception to be expected return and some measure of risk. The thrust of these models is to show that the observable phenomenon of diversification is reconcilable with the pursuit of an optimal portfolio. As Harry Markowitz [5, 6] has led the way, it becomes a matter of quadratic programming to find the portfolio which gives the highest expected return for any specified level of risk. Expected return for a portfolio is the weighted average of the expected returns of the portfolio's component assets. Risk is measured by the portfolio's variance of expected returns and is computed from the variances of the component assets as well as the covariances between them.

1. Up to this point, the terms "perception," "discrimination" and "evaluation" have been used interchangeably. This implies the belief that at least as far as investments are concerned, there is no important difference between a person's perceptual space and his evaluative or preference space. As Green and Carmone [4] have cautioned, however, this assumption cannot be safely made for attitude research of most stimuli. Strictly speaking, the current study is one of investment and portfolio evaluation. This is the interesting issue and is forced by the nature of the data used in the present study. Nonetheless, the terms referring to "dimensions," "factors" or "variables" of "preference," "perception," "discrimination" or "evaluation" will be used interchangeably throughout this report and will be taken to be the same thing as the notion of "investment objectives."

The set of highest-return portfolios over the range of risk defines the familiar efficient frontier. The selection of which one of these portfolios is best for a given investor depends upon his individual weightings of the risk and return arguments in his utility function.

Markowitz's landmark study has been carried forward in many directions by numerous investigators. Some, such as Sharpe [7], have investigated ways in which the quadratic programming problem can be simplified to facilitate the model's use as an operational tool.

Other investigators have explored statistics other than the variance (or standard deviation) of the probability distribution of expected returns as possible measures of risk. Such studies are often stated to be motivated by "the desire to link theory with practice." The need is felt to modify the theoretical model to make it more "realistic" in terms of intuitively appealing notions of risk. The question has been asked: How realistic is it to assume that risk-averse investors will regard a given positive dispersion about the expected return value as being equally undesirable as an equal but negative dispersion about the same mean? Thus, investigators have proposed model modifications using statistics such as semi-variance and partial variance.¹

Criticism can be leveled at any proposal for modifying the Markowitz model whose aim is to better reflect "reality" but which is

1. For example, see Ilano [8] and Pinches and Kinney [9]. Markowitz also suggested the possible use of semi-variance [5, 6].

not rooted in empirical evidence. Models of this sort represent an unsatisfactory compromise of normative and positive aims. If the intent is to establish a true description of investment selection, a better research strategy would be to start from a position, and use a technique of empirical inference, which does not require the prior acceptance of certain tenuous assumptions which are basic to the Markowitz model; namely, an optimizing objective on behalf of investors and their conscious evaluation of a portfolio's risk. Serious observers of actual investor behavior have raised issues which contest both of these premises of the Markowitz model. The following observations by Smith are relevant:

"Conversation and correspondence with investment managers of various types of institutions lead to several hypotheses about current selection procedures. First, managers appear willing to satisfice rather than optimize in many decisions. That is, they choose the first security that satisfies a particular set of specifications instead of searching for the single security that best achieves the specifications. Second, selection decisions are made on a security-by-security basis rather than by focusing on the total portfolio as a distinct entity with measurable characteristics. Third, a frequent practice is to invest equal sums in the securities chosen for the portfolio. Fourth, managers attempt to accomplish diversification simply by investing in a large number of securities, but without any consideration of the interrelationships between securities. Fifth, although expectations for an individual security typically are not made explicit, the orientation of many investment managers appears to be on the 'downside' of an implicit probability distribution for that security..." [10]

We saw that covariances (intercorrelations) play the major role in providing Markowitz's normative rationale for diversification. However, one of Smith's key observations is that investments are selected

on a security-by-security basis without regard to their interrelations. With diversification being a real feature of investor behavior, this implies a need to revise the normative view of investor's use of risk in portfolio evaluation.

Several explanations of diversification are possible which do not require the assumption of a conscious regard of covariation. For example, the perceived risk, R_i , of an investment i , may be described in a strictly positive sense as being directly proportional to the amount the investor already holds in his portfolio. That is, marginal utility for risk may be a negative linear function of h_i , the percent of the investor's portfolio already represented by investment i ; namely -

$$\frac{dR_i}{dh_i} = -k_i h_i$$

On integrating (with the integration constant set equal to zero), we get:

$$R_i = \frac{-k_i}{2} h_i^2 = -g_i h_i^2$$

Here g_i may be regarded as the investor's risk aversion coefficient for investment i . Still further, g_i may be decomposed into $b\rho_i$ where ρ_i is the risk value of the investment on some risk attribute scale as perceived the same by all investors. The coefficient, b , represents the investor's individual risk weighting coefficient in his utility function. The portfolio selection process at any instant for investor S is then:

$$\text{Maximize: } U_{iS} = a_S \sum_i \mu_i h_{iS} - b_S \sum_i \rho_i h_{iS}^2$$

$$\text{Subject to: } \sum h_i = 1.0 \quad , \quad h_i \geq 0.$$

where U_{iS} stands for investor S's total utility for asset i, μ_i is the expected return of a unit of asset i, and a_S is the investor's return weighting coefficient.

This quadratic programming model can lead to optimal portfolios which are diversified. Any investor, S, can be uniquely represented as a point or vector in a common evaluative space in terms of his characteristic return and risk weighting coefficients, a_S and b_S , respectively.

The foregoing model is not being suggested as a description of actual behavior. Indeed, as Smith observed (see also Clarkson, page 17), there is reason to doubt the entire notion of investors acting to optimize utility. The only point being made here--one that is crucial to this study--is that the phenomenon of diversification does not require the assumption that investors consciously seek investments which are dissimilar as to their risk features in the sense of having low inter-correlations. Indeed, explicit consideration of risk in any form may be a weak factor in explaining which companies and industries are selected for a given portfolio if, for no other reason, the difficulty of discerning risk differences. Some or most observed diversification

may be related to considerations other than direct perceptions of volatility; namely -

- 1) Institutional constraints on the amount of a given security that a mutual fund can hold. The Investment Company Act of 1940 requires that at least 75 percent of a mutual fund's portfolio must be "diversified" - meaning that no security holding within this 75 percent (other than U. S. government securities), may represent over 5 percent of the fund's total assets nor over 10 percent of the outstanding voting securities of the issuer. The Internal Revenue Service Code also imposes similar constraints.
- 2) Dynamic evolution of information affecting expectations of return. This, coupled with transaction cost of making portfolio changes, would tend to direct new investment dollars to new securities without the simultaneous switch of an investor's current holdings.

No matter what are all the exact reasons for diversification, if we can at least say that the objective of minimizing intercorrelations of price movements is not an important effect in practice, then it is reasonable to argue that an investor will form a portfolio of several or many assets which are largely similar in terms of the investor's investment objective weightings on the basic psychological attribute dimensions of evaluation.

Another point of divergence between positive and normative models of investment behavior is the expected return dimension of preference. Observers of actual behavior call attention to sharp distinctions that investors frequently make between growth and income objectives. This is evidenced in the Growth/Income basis for most

formal mutual fund classifying systems. The meaning of this distinction of objectives is that investors may regard a dollar of capital appreciation differently than a dollar of dividend or interest. In part, this reflects differences in tax treatment and an investor's differential weighting of these dimensions reflects his particular tax status. However, at a deeper level, the separation of perceptual dimensions of growth and income may also reflect the notion that return as capital appreciation generally is less certain than return in the form of income. In other words, the notion of risk or uncertainty may not be a distinct dimension of perception but may be embedded in and tend to help define the psychological factors associated with the terms "growth" and "income."

* * *

The intent of this section has been to call attention to some of the challenges that have been made to the normative Markowitz type model as an approximation of how people really behave. If an adequate positive description of investor behavior is the objective, the research should not proceed by trying to rebuild the conventional normative model with its several vulnerable assumptions. A better research strategy would be to develop a positive model with as pure an empirical inference technique as possible and then at the end attempt to reconcile the results with predictions of normative theory. In this light, what is needed is an empirical research plan and technique which can elicit

the factors of investment preference from data on actual portfolios while requiring as few prior assumptions as possible. Next are considered some past strategies of empirical research which have been used to study investor portfolio selection behavior.

HISTORIC EMPIRICAL INVESTIGATIONS

Implicit in the foregoing sections is the idea that the procedure of portfolio construction is a two-stage process; namely -

- 1) Discriminating between each of the candidate investments in terms of the basic dimensions of preference (utility) and, then,
- 2) Selecting among the population of eligible investments through some decision process to construct an optimal or satisfactory portfolio.

Most empirical research concerning portfolio construction has as its focus the second of the two steps above. The research aim is to find an adequate representation of an investor's decision mechanism for portfolio construction rather than the more primal quest of uncovering an investor's true frames of reference for distinguishing among different investments and portfolios.¹

1. In presenting his landmark model Markowitz was careful to draw attention to which of the two issues his normative model was addressed. He stated: "The process of selecting a portfolio may be divided into two stages. The first 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 performance and ends with the choice of portfolio. This paper is concerned with the second stage ... We have not considered the first stage, the formation of relevant beliefs on the basis of observation." [11]

The pattern of empirical testing of models of decision-making is consistent. Dimensions of utility are assumed. A decision mechanism hypothesized. Predictions are made with the model. The predictions are compared against actual results or known behavior. Confirmation of the model is made if the actual results do not differ in a "statistically significant" sense, if the model "performs better" than a mechanism which generates results by random or otherwise naive procedures, or simply if familiar visual patterns in real observations such as clusters, appear in the model's output.

Portfolio models which have been tested by this empirical inference approach have been both normative and positive varieties. Following is a brief description of an example of each, both of historical significance and each having important implications for the present study.

Geoffrey Clarkson

Clarkson's dissertation [12] represents a positive model of the portfolio selection process. His approach was to design a computerized simulation model of the deliberative processes of a practicing bank trust officer which could predict the investment selections that the officer would actually make over a range of conditions.

Clarkson's first step was to classify each of the bank's trust accounts into one of four categories--Growth, Growth-Income, Income-

Growth, and Income--based upon information concerning the respective investor's stated objectives and his financial and legal status. His second step involved gathering financial data for each company and industry whose securities appeared on a master list of investments approved for bank trusts. Then, the hypothesized decision model was used to select a portfolio for each trust account through a process of comparing the company and industry characteristics of individual securities against the account's investment objectives.

The decision mechanism did not derive from any normative economic theory. Rather it was founded in Clarkson's careful observations of the heuristics--the rules and guidelines--employed by the human trust account officer. Thus, for example, portfolio diversification resulted from the guideline that all accounts participate in at least five industries.

The investment selection model itself involved a recognition structure called a "discrimination net" and a stepwise matching process to reach decisions. These "nets" to screen candidates for acceptance or rejection comprised a hierarchy of tests where the lower the level, the more detailed was the measure of the company's performance considered.

The simulation model accurately reproduced the portfolio selections of the human trust manager. Clarkson was circumspect to emphasize, however, that the high correspondence between predicted and

actual behavior tended to "corroborate" the structural elements and assumptions of his model but could not be interpreted as having "verified" them.

Some implications of Clarkson's study of actual portfolio selection behavior and corroboration through model testing are important to the present study. First, his study did not support the hypothesis that investors make a conscious effort to evaluate risk but that the notion of risk is implicitly embedded in the decision mechanism such as the guidelines for diversification.

Second, the idea that investors attempt to optimize utility in portfolio selection must be abandoned in favor of the notion of satisficing. An optimization hypothesis implies that an investor examines all available alternatives before making his selection. Instead, the decision process is sequential--investors make choices by considering alternatives one at a time and by accepting the first one that passes the series of tests. This means that the order of processing is important and that the absence of an investment from a portfolio does not necessarily mean that the candidate was evaluated and dismissed. It may mean that the investor reached his budget limitation before the security could be appraised.

Finally, both company characteristics and industry characteristics are used by investors in evaluating individual investment opportunities. Investors associate specific industries with different investment objectives.

Donald Farrar

Farrar's objective was to subject the Markowitz theoretical model to an empirical test. He asked whether the normative model could be shown to reflect actual portfolio selection behavior in the sense of accounting for observed differences between the holdings of real mutual funds. Do the funds in fact provide lower risk for given returns or higher returns for given risk?

Farrar started with the suggestion that in building their portfolios, mutual funds strive to maximize the following utility function:

$$U = \mu - A\sigma^2$$

where A is the risk-aversion coefficient, whose value characterizes each fund. The objective function in this form is equivalent to the Markowitz programming goal of minimizing a portfolio's variance, σ^2 , for any level of expected return, μ .

Farrar selected a sample of 23 mutual funds and initially classified each fund's holdings into 51 industry groups for common stocks and 7 groups for preferred stock, bonds and cash. Next he computed the intercorrelations between these 58 industry and asset categories from the covariations in their respective historic price movements over a ten-year period. Then he factor analyzed the correlation matrix to get a smaller set of distinct asset groups, each one comprised of one or more of the 58 original categories. Eleven

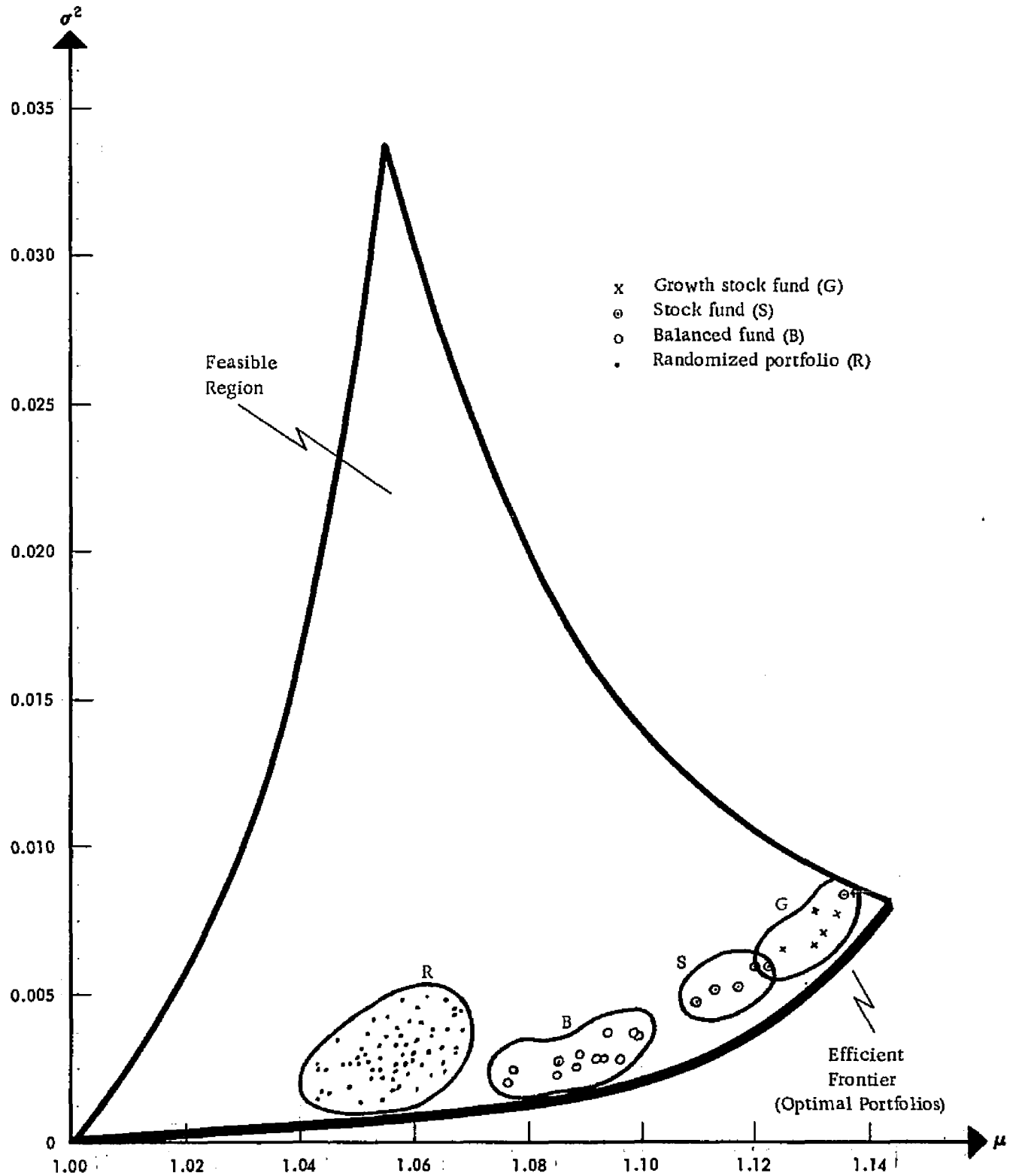
new asset categories were found, corresponding to the extracted factors, and the holdings of each portfolio were redefined in terms of these new, more basic variables.

For estimates of the expected returns and risk for the next one-year period for each of the eleven basic groupings, Farrar took the simple projection of the most recent trend, taking as the appropriate trend pattern for each grouping the one which would have provided the best least squares fit if it had been used as a predictive mechanism throughout the ten-year period covered. With these estimates of μ and σ^2 for each of the groupings of securities, and by varying the value of the risk aversion coefficient, A , he used the Markowitz model to generate a series of optimal portfolios forming a locus as the efficient frontier. The μ and σ^2 for each optimal portfolio as a whole was calculated from the expected returns and variances and covariance, respectively, of the portfolio's component assets. Similarly, for each of the 23 actual portfolios a μ and σ^2 were computed.

At this point Farrar plotted both actual and optimal portfolios in the same μ, σ^2 space. He argued that should the actual portfolios form a swarm of points which hover close to the efficient frontier, it could be interpreted as lending credence to the idea that the Markowitz programming model is a description of the portfolio decision policies of mutual funds.

Figure 1

FARRAR'S MEAN-VARIANCE FEASIBLE REGION
SHOWING ACTUAL, OPTIMAL AND RANDOMIZED PORTFOLIO'S [13]



The actual fund portfolios did fall close to the locus of optimal portfolios. Even more encouraging to his thesis, Farrar said, was the tendency of funds to cluster along the boundary in almost perfect groups according to their prior known stated investment objectives (see the G, S and B fund clusters in Figure 1).

Farrar claimed that, to the extent that actual portfolios occurred in the same region as optimal portfolios, support was lent to the idea that the Markowitz decision mechanism was that used in practice by mutual funds. By nature of the assumptions upon which the Markowitz model is built, the implication is that μ and σ^2 are the real perceptual dimensions of portfolio discrimination and preference.

A NEW STRATEGY FOR RESEARCH

Consideration of the two empirical studies above prompts a relevant and vexing question: How well can any test of a portfolio selection model confirm the perceptual dimension assumptions which underlie the model? Much controversy and philosophic debate concerning the scientific method has questioned the degree to which the validity of the assumptions and structure of any model are confirmed by the model's ability to make correct predictions. If disagreement between actual and predicted results are observed, then we must infer that the model is an inadequate representation of the world. However, if agreement is observed, the most we can say about the micro-level features of the model is that they have been confirmed only in the sense

of not being refuted by the facts. We cannot automatically be sure that a model with different assumptions or structure will not make predictions which are more accurate. Before we can comment on the importance or correctness of any model assumptions or elements, we have to test the sensitivity of the model's performance to changes in these model features.

For example, it is impossible to draw the conclusion from Farrar's study that μ and particularly σ^2 have been confirmed as the true dimensions of investment evaluation.¹ Farrar "generated" 100 randomly diversified portfolios and showed that they, like the actual portfolios managed by mutual funds, fell along the efficient frontier (see the R fund cluster in Figure 1). The informational content of the σ^2 dimension appears negligible as an explanation of portfolio differences, especially with the efficient frontier as the datum.² Indeed, it appears that the observed clustering can be explained by the μ dimension alone.

1. Farrar's work is not being faulted in these or other remarks in this report. It is emphasized that Farrar's research objective is related only tangentially to the research purpose of the present study.

2. This effect would give a logical basis to the belief that portfolio managers make no serious attempt to appraise the covariation between selections. The simple rule of restricting the percent of total assets held in any one security would seem to give an adequate probability of a low portfolio variance. To test this, Fisher and Lorie [14] investigated the dispersion of returns for randomly selected, unmanaged portfolios consisting of specified numbers of stocks, ranging from one to 128 issues. Their results showed a very sharp reduction in portfolio risk, as measured by the degree of dispersion of sample returns compared to the general market, with a portfolio size of as few as eight random stocks.

In light of everything, a preferred research strategy for exploring the dimensions of portfolio perception would be one that is not confounded with the issue of establishing the correct decision process. It would be a research design requiring as few prior assumptions as possible. It would be inductive logic--empirical inference using actual data--to discern the dimensions or underlying factors which account for observed differences in real portfolios. These requirements are met by and tend to define a factor analytic-type approach.

Factor Analysis

At the core of factor analysis is the notion that measurable similarities between items being compared can be traced to underlying effects common to each. Each item, or "subject," can be described in terms of a linear compound of the various common effects, idiosyncratically weighted, plus a unique effect. Each subject can be characterized in terms of his individual weightings of, or "loadings" on, the common factors. The conditions of this general factor model are the same as the features of the model of investment perception discussed on page 7.

Factor analysis as a general technique starts with subjects represented as points or vectors in an n-dimensional "test" or "stimuli" space. Each subject point or vector is positioned according to the subject's respective scores or responses on the n tests. Correlations (co-variances) between the responses of different subjects on the tests are

regarded as manifestations of the common factors acting on or within each subject. Classical factor analysis transforms the stimuli space into one of prescribed lower dimensionality such that the new configuration of subjects best reproduces the subjects' original intercorrelations. The dimensions of the reduced space are the extracted factors which account for differences between subjects in the original stimuli response data.

Once a set of factors is found, their interpretation is the next step. The degree of "meaningfulness" of a factor solution is a controversial issue. Depending on the nature of the problem, elicited factors may be regarded anywhere over a bipolar continuum ranging from interpretable verities of causation and explanation to merely mathematical constructs of classification and description.

In the first instance, the belief is that the factors are the causes of the observed patterns in the data. Thus, an intelligence quotient factor uncovered by factor analysis might be considered a "cause" of individual differences in stimuli responses. Similarly, a market-correlation coefficient of security price movement as revealed by factor analysis¹ may be treated as a real cause of price movements from which predictions can be made. Often when factors are regarded in this causal sense, the research aim of factor analysis is as a tool

1. See studies by Feeney and Hester [15], King [16] and Meyers [17].

for hypothesis testing. A priori hypotheses are set forth concerning dimensions of personality, behavior or perception and factor analysis can be performed to test for their empirical existence.

In many instances researchers are unwilling to go as far as to credit reality to their discovered factors. The second point of view of revealed factors is that they are merely mathematical concepts which serve to best reproduce the variations observed in a set of data. Factor analysis becomes a method of analyzing data to determine whether the variations can be accounted for adequately by a smaller number of basic categories or variables than when the investigation was started.

Rummel [18] has distinguished between a number of related research aims associated with using factor analysis in a descriptive/ classifactory sense including: 1) parsimony or data reduction, 2) categorization, 3) scaling, and 4) data transformation. In the area portfolios, at the individual investment level, Farrar's use of factor analysis is an example of the technique's use in a descriptive, data reduction and transformation sense. Each mutual fund could be thought of as being a subject defined in a stimuli space where the stimuli are the universe of securities eligible to be held. A fund's holding in any security, say expressed as the percent it represents of the fund's total portfolio assets, could be regarded as the fund's response to that particular security stimuli test.

As discussed earlier, Farrar's objective was to portray his 23 fund-subjects in a reduce stimuli space of manageable proportions. The population of securities eligible to be held represented perhaps several thousand variables. Despite its obvious shortcomings, Farrar said, the absence of a practicable alternative served to justify aggregation of the individual company securities into his 58 initial industry and asset groupings.

Continuing, Farrar felt the need to reduce the 58 stimuli space still further into a smaller set of factors corresponding to groups of industries and assets which are "meaningfully" diversified. The following passage by Farrar provides his argument for using factor analysis to extract from the correlation matrix of the 58 investment groups a smaller set of 11 more basic groups. It is an important passage and is included here as a convenient bridge to an explanation of the use of multidimensional scaling in this report. Farrar said:

"Suppose that comparison is desired between two portfolios, each of which contains three alternative assets A_1 , A_2 and A_3 . Suppose further, that intercorrelations between the aspect of each asset's behavior which is considered to be relevant (for example, its price, yield, etc.) is represented by the matrix of correlation coefficients

A_1	1.00		
A_2	.99	1.00	
A_3	.02	-.03	1.00
	A_1	A_2	A_3

Should one portfolio contain a high concentration of A_1 and the other of A_3 , an ordinary, least squares goodness of fit test, for example,

$$D = \sqrt{\frac{1}{n} \sum (h_{1i} - h_{2i})^2}$$

where h_{1i} is the proportion of the first portfolio which is held in the form of asset i , $i = 1, 2, \dots, n$; h_{2i} is the proportion of the second portfolio that consists of asset i , and D is the measure of goodness of fit, concludes, quite correctly, that the portfolios are markedly different. They would not, at least, be expected to behave with much similarity over time.

"Should the two portfolios be concentrated in securities A_1 and A_2 , however, the preceding test would indicate an equal degree of dissimilarity. In this instance, unfortunately the conclusion's validity is much less obvious. When assets A_1 and A_2 share a correlation coefficient of .99, it is not at all clear that diversification between them is meaningful. Replacing the one by the other will, ordinarily, not affect the portfolio's behavior in any but the most minor of ways." [19]

"Meaningful diversification," as Farrar used the term above, can be interpreted to be equivalent to the notion of being substantially different with regard to investors' dimensions of evaluation. What Farrar has suggested is that comparing portfolios by differences in their holdings, by itself, does not necessarily reveal anything about differences in their investment objectives. We can only establish the meaningful diversity between investments and the portfolios they comprise in terms of the investments' pairwise covariations with regard to pre-determined "relevant" behavior variables, e. g., price, yield and so forth.

Farrar's study was an empirical test of an investment selection decision model and implicitly embodied prior assumptions concerning the relevant factors of investment perception and preference. Unlike Farrar's investigation, the present study is a search for the relevant dimensions of preference and, consequently, seeks to avoid making such prior assumptions. The purpose is to find a factor structure which explains differences in fund holdings not to test a hypothesized one, one that is implicitly embedded in a decision-making model.

The essence of Farrar's remarks above is that comparing portfolios on the basis of their holdings alone would not provide adequate clues if our research aim is to uncover the dimensions of preference. And yet, intuition suggests that there is a great deal of informational content in the knowledge alone of the industry and asset groupings of the portfolios of a large sample of mutual funds. Perhaps there is sufficient information from which to envince factors which account for the differences in holdings--that is, factors which relate to differences in investment objectives of various funds. The remainder of this report aims to present a method by which this problem now can be solved. The solution takes advantage of the new technology of non-metric multidimensional scaling, a method that only become available in an operational form within the last ten years, subsequent to Farrar's remarks. Inasmuch as the methodology is quite new and likely to be unfamiliar to many students of finance, the next sections

give an introduction to the conceptual foundation of multidimensional scaling in the context of the investment evaluation problem.

INVESTMENT SIMILARITY

Assume the existence of dimensions of investment preference along which each investment--say, represented by its respective industry group--can be scaled. If no diversification mechanism existed, each investor's "portfolio" would consist of a single industry group that he judged to be best. However, with a diversification process present, but assuming no or weak interdependency between investment choices, a portfolio should contain a selection of investments which are similar in terms of their scale values on the factors of preferences. A comparison of the portfolio holdings of any given pair of mutual funds which have sharply different investment objectives (weightings on the dimensions of preference) should show that they hold few industry groups in common.

Because of the constraint of an investor's budget and because investments are evaluated sequentially, not all of those investments that are similar to an investor's "best" investment will be selected. Thus, for any given portfolio the exclusion of any industry group will not necessarily imply that it was perceived to be dissimilar to those industry group investments that were selected. For this reason, a comparison of the portfolio holdings of any given pair of mutual funds which have the same investment objectives still could find that they

hold few industry groups in common. Nonetheless, for a large sample of portfolios, similar industry groups would tend to occur together more frequently in the pairwise comparisons than those that are judged not to be similar. If the individual industry groups are viewed as points in a common perceptual space, and are positioned as to their perceived differences across all investors, then it should be expected that the closeness between any pair of investment points should relate in some manner to their degree of co-occurrence in a sample of real portfolios. The larger the sample, the more likely are the co-occurrence measurements to accurately reflect the "proximities" among the true configuration of industry group points in the common perceptual space.

Going further, each portfolio also could be portrayed or "mapped" as a unique point in the same industry group perceptual space. That is, a portfolio is no more than a composite investment. It would be positioned in the investment perceptual space such that its distance from each industry group which it holds would be in proportion to the percent of the portfolio's total assets that the group represents. Again, because of the budget constraint, industry groups which are not held in the portfolio are unreliable indicators of investment preference and would not be given weight in positioning the portfolio point. At the extreme, a single-asset portfolio would be a point coincident with the industry group point of which it is uniquely composed. Clearly,

portfolios which are comprised of a relatively high proportion of "similar" industry groups will be positioned close together.

Two key notions in the foregoing are: 1) co-occurrence of industry groups in a sample of portfolios is indicative of the perceptual similarity between the groups, and 2) similarity can be given a spatial interpretation. The proximity or closeness of any pair of subjects, be they portfolios or individual industry groups, depends upon their scale values along the attribute dimensions which define the perceptual space. If we knew our reference axes and had coordinate values for each portfolio and our research question concerned the degree of similarity between portfolios, it is clear we could compute their interpoint distances as a measure of their similarities.¹ But the transpose question is more interesting: Given measurements between all pairs of funds on some index of dissimilarity (or similarity) can we regard these measures as interpoint distances (or complements of distances) in the perceptual space and reconstruct the configuration of points? If the answer is yes, will such a configuration of points be unique? If we cannot assume that the computed dissimilarities are perceptual distances but that they are related to distances, say, according to some monotonic rank order function, can a configuration solution still be obtained? That positive answers can be given to all these questions is a result of the development of nonmetric multidimensional scaling.

1. More correctly, their dissimilarities--a zero distance between points means identical items, hence zero dissimilarity

MULTIDIMENSIONAL SCALING CONCEPTS

In two pioneering articles in 1962, Roger N. Shepard [20] ushered in the technique of nonmetric multidimensional scaling. Using ordinal scale (rank order) dissimilarity data alone, his algorithm was able to give a metric scale configuration of subject points positioned in a space of prespecified dimensionality so that the ranks of their inter-point distances were highly correlated with the ranks of the pairwise dissimilarity input data. He showed that if the number of subjects were sufficiently large relative to the dimensionality of the configuration space, rank order data are sufficient to yield essentially unique metric solutions. By being able to accept rank order data, the assumption need not be made that the dissimilarity data are linear transformations of distances in the evaluative space. The only assumption required is that dissimilarity (or similarity) be a monotonic function of distance. Indeed, the output of the Shepard procedure itself provides a clue to the form of the function relating dissimilarity values to distances.

A number of computer programs have been, and are continuing to be developed which are improvements upon Shepard's initial work. For example, the computer program M-D-SCAL will take as input data a matrix of similarity or dissimilarity values between a set of subjects. The option exists to treat these values as metric (interval) scaled distances (distance is a linear transformation of dissimilarity)

or simply as rank order values (nonmetric multidimensional scaling).¹ The program will produce a configuration of subject points and the distances between them will be ratio-scaled. The output configuration is unique up to a similarity transformation; that is, the configuration can be translated, rotated, reflected, magnified or compressed without changing the rank order of the interpoint Euclidean distances. This means that there is no unique orientation for a set of reference axes in the solution space. The projections of the subject points onto any set of chosen axes will be interval-scaled. In short, by making maximum use of the informational content of a large quantity of weak nonmetric scaled input data, multidimensional scaling is capable of upgrading the information to a stronger metric scaled output.

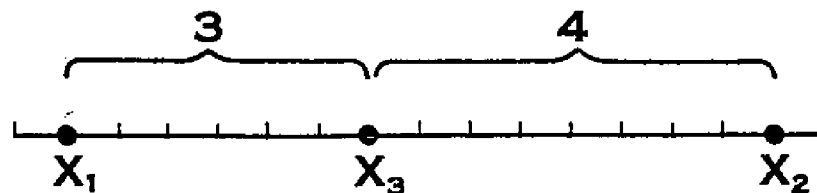
Like factor analysis, multidimensional scaling does not provide identifications for the dimensions; labeling of reference coordinates follows after the scaling step. Moreover, the correct number of factors is normally not known ahead of time and configuration solutions can be called for over a prespecified range of dimensionalities, say, one to five. For each solution space a measure of badness-of-fit, called "stress," is given by the scaling program to indicate the degree of mis-match between the interpoint distances in the derived output configuration and the corresponding dissimilarity input data.

1. Frequently, but not necessarily, distances will be specified as Euclidean. This point is discussed in Chapter 5.

Stress measurements play an important role in determining the appropriate number of dimensions for a given scaling problem. To see this, assume we have three subjects, say mutual funds X_1 , X_2 and X_3 , and have judged or computed in some manner the following metric scaled values of dissimilarity between them:

X_1	0.		
X_2	7.	0.	
X_3	3.	4.	0.
	X_1	X_2	X_3

These dissimilarity values would be inputs to a scaling program and the one dimensional solution output would be:

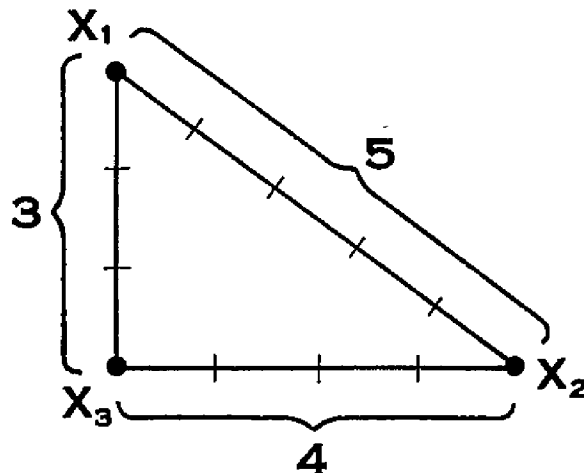


Interpoint distances within this configuration exactly match the input dissimilarities (up to a proportionality transformation) thereby indicating that one factor alone is sufficient to account for differences between the subjects.

However, now assume that the observed dissimilarities are the following:

X_1	0.		
X_2	5.	0.	
X_3	3.	4.	0.
	X_1	X_2	X_3

No placement of subject points along a one-dimensional scale can exactly reproduce these dissimilarities as Euclidean distances. Any unidimensional solution will have some error reflected as stress. Instead, to obtain a perfect fit the subjects must be arranged as:



Two dimensions are needed for a perfect spatial representation imply that two factors are at work contributing to variations observed in the data.

From this it would appear, therefore, that the search for the dimensionality of the correct scaling solution involves generating solutions of higher and higher dimensionality until the first one with

zero stress was found. However, except for synthetic data, perfect zero stress configurations almost never will be reached.¹ Another determinant of the stress value for a configuration is measurement error or noise in the original data which may manifest as intransitivities.²

Nonetheless, even for noisy data, with a large number of subjects, the derived configuration should well reflect the "true" configuration of subjects in perceptual space. As the number of subjects, n , increases the number of rank order constraints increases almost as $1/2 n (n-1)$. However, to position a set of points in r dimensions, we need only approximately rn numbers corresponding to the coordinates of each point on each dimension. For example, if $r = 2$ we need only 100 pieces of data to fix 50 points. In contrast we have 1225 pieces of data with all dissimilarity or similarity measurements. In short, a larger value of stress can be tolerated as the number of subjects increases.

1. Except until the trivial case where the number of dimensions is one less than the number of subjects if the input dissimilarities are treated as metric distances or two less than the number of subjects if the input data are treated as rank order distances; see Wagenaar and Padmos [21].

2. Intransitivities are violations of the triangle inequality axiom of distance in a metric space. This axiom says the distance from point x to point y must be less than or equal to the distance from x to y indirectly through point z , or

$$d(x, y) \leq d(x, z) + d(z, y)$$

It is these inequality (rank order) constraints which tend to fix the configuration of subject points in the scaling solution.

RESEARCH USING MULTIDIMENSIONAL SCALING

At this point it should be clear that the research spirit of multidimensional scaling is like that of factor analysis. The aim of both is to reveal latent factors which account for observed variations in a set of data. The factors uncovered may be regarded in a causative / explanatory sense or in a classificatory / descriptive sense or, somewhere in between.

Just as factor analysis and multidimensional scaling are in the same research spirit, so are they alike as technical procedures. Both methods assume an additive model of perception or evaluation. With both methods the concern is the discovery of the frame of reference using a configuration of points or vectors merely to get the particular reference axes. In classical factor analysis the geometric configuration is a spread of vectors, each one representing a subject, the pattern of vectors being a function of the correlations among subjects' responses to various tests or stimuli.

Multidimensional scaling differs from factor analysis primarily in that the pattern of subject points is related to the more general notion of proximity--similarity or dissimilarity--as a distance. However, factor analysis and multidimensional scaling are really not to be contrasted. Correlation coefficient can be considered to be a type of similarity measure and, as can be shown, distance can be equated to a function of scalar products between subject vectors. Therefore, a

factor analysis could actually be accomplished using nonmetric multidimensional scaling although inefficiently.¹

Many measures of proximity between subjects are possible. For convenience, three general classes can be distinguished. First and most often, similarities or dissimilarities are obtained as direct judgments by human respondents to pairwise comparisons of stimuli. Generally, it is possible for human respondents to express their judgments only in rank order terms. Therefore, because nonmetric scaling is the only way of adequately handling data in such weak form, these techniques are particularly suited for perception and attitude research. Understandably, the large majority of applications of multidimensional scaling have been in the fields of marketing and mathematical psychology.

The second way of obtaining proximities is to use some single aspect of how people respond to pairs of test stimuli as a surrogate for direct judgments of stimuli similarity. For example, the frequency in which two stimuli are confused by respondents or, the average time it takes for human subjects to discriminate between the stimuli, can be taken as behavioral measures of stimuli similarity or dissimilarity.

The third way of developing proximity values between pairs of subjects is to derive them by computation over the subjects' profiles of

1. See Shepard [22]; also Tucker [23].

responses to a standard set of stimuli. The stimuli do not necessarily have to be behavioral variables and, indeed, the subjects need not be people. For example, one scaling study has been reported [24] where the "subjects" were archeological sites and their similarities were computed over several cultural and physical variables. Predictably, the choice of test stimuli will be an important determinant of the type and meaningfulness of results to be obtained from a scaling experiment. Choosing a wrong stimuli set can impose a factor solution, one which is not related to the original research question.

The present study of portfolio evaluation involves proximity measures of the last type described above, derived similarities. A sample of mutual funds are the subjects; industry and asset group classifications of securities are the test stimuli; the percentage of a fund's total assets that it has invested in an industry group is the fund's "response" to that stimuli. Thus, a mutual fund's portfolio can be regarded as its profile of stimuli responses and, in this way, readily available "objective" data on portfolio holdings can be used.¹

1. The stated purpose of the present study is to reveal the preference structure of portfolio investors and to scale existing portfolios with respect to this preference structure. The use of mutual fund portfolios as the basic data for scaling will lead to a factor solution which reflects the preference structure of those who build the portfolios, namely, the fund's portfolio managers and committees. Thus, the solution will not derive directly from the expressed preferences of mutual fund investors themselves. Nonetheless, the assumption is made unreluctantly that the dimensions of preference revealed in this study are relevant to portfolio investors. It is only reasonable to believe that mutual fund managers share the same perceptual framework as their investing public, that they are responsive to their clients' preferences and motivated accordingly in selecting their fund's holdings.

This avoids the direct solicitation from actual portfolio investors of their evaluation opinions. The use of a portfolio holdings stems from the belief that the best manifestation of an investor's preferences for different investments is the extent to which he actually allocates his resources to them. In short, what an investor does is a more reliable measure of preference than what he says he might do.

Many possible formulas for transforming profile data into proximities have been suggested. Correlation coefficient qualifies as one index. Another possibility is Farrar's measure of goodness of fit discussed on page 29. Consideration of this formula reveals that it is essentially equivalent to the Pythagorean theorem for computing Euclidean distance between portfolio subjects in Farrar's original 58 industry and asset group stimuli space. This fact, as well as his emphasis on the idea of degree of portfolio dissimilarity further serves to demonstrate how close Farrar came to anticipating a spatial proximity notion of portfolio differences.

The answer to the question, "which measure for computing similarities is correct," must be based on the features of the particular problem. And the choice is crucial. Chapters 2 and especially 5 explain the index designed for the present study.

A FINAL PERSPECTIVE

This introduction has covered a lot of ground. First, it explained the need to provide mutual fund investors an improved comparison of

the current investment qualities of available portfolios. It is a call for a method of displaying the portfolios of mutual funds within a metric space spanned by coordinates that correspond to investors' multi-attribute dimensions of choice. Next, it showed that nonmetric multidimensional scaling is a new statistical technology that leads to metric multidimensional scales which account for variations in a set of observed data. In short, the introduction has been an attempt to demonstrate the nexus between a problem and a technique of analysis.

At the outset it must be made clear that nonmetric multidimensional scaling is still in its infancy. Any study at this stage which employs multidimensional scaling methods and which suggests that they can be applied in a routine fashion, denies the truth that the choice of correct procedure is highly sensitive to the problem context. Results are dependent upon which scaling algorithm is used, which concept of proximity is employed and so forth. The theory resolving these issues of experimental design and including the question of statistical significance has not yet been established; aspects of the research must proceed on a trial and error basis.

The use of multidimensional scaling in this study is in a somewhat unusual quasi-behavioral context relying on non-subjective data of investor (mutual fund) behavior profiles (portfolio industry holdings). In several important respects the study is a distinct departure from empirical work that has been reported heretofore in the literature both

from the standpoint of the theory of investments and from the standpoint of multidimensional scaling applications.¹ This suggests that the study has been motivated by dual objectives:

- 1) To develop a method for empirically displaying the factor structure of portfolio differences, and
- 2) To explore the performance of multidimensional tools under somewhat new conditions and data.

With respect to both, this study only claims to be a beginning. And as a first effort, the intended objective is not as ambitious as to find a factor structure which explains variations in the data but, more modestly, to reveal a structure which describes the variations. Thus, this report is not presented as analysis of investor behavior in the sense of psychometrics. It is to be viewed essentially as a descriptive statistical study of the variation in a set of data, the portfolio holdings of a sample of mutual funds. Many avenues to explore the psychological implications of the study results remain open for further investigation.

The remainder of the report comprises five chapters. First, the basic experiment design is introduced with a discussion of data selection

1. The only known application of multidimensional scaling to the problem of investment evaluation in the literature is a study by Green and Maheshwari [25]. Their study is unlike the present investigation in several major ways. It was at the level of selecting individual investments, not portfolios. It was a test of a normative model of investment perception, not a search for a positive model. It was based upon directly judged investment similarities by a panel of investors (graduate students), not based upon derived similarities from the actual investment selections of real investors.

and preparation. The derivation of a matrix of similarity measures between mutual fund portfolios is traced from the initial selection of mutual fund subjects through the aggregation of each subject's holdings into standard investment groups.

Next, the key actual multidimensional scaling experiment is explained. The input data are the computed mutual fund portfolio similarities representing a single point in time. The resulting configuration of fund points is presented and an underlying factor structure is revealed and interpreted using a variety of techniques including cluster analysis and multiple regression.

Then follows Chapter 4 which explores the responsiveness of scaling results to changes over time. Data preparation and similarity scaling steps are repeated for a total of four points in time. Methods are investigated for measuring the congruence of the chronologically different solution configurations.

Chapter 5 describes the sensitivity of experimental results to several important methodological variations in data preparation and scaling design. This chapter provides much of the logical underpinning for the key experiment design presented in Chapters 2, 3 and 4.

Finally, there is a concise summary of the implications of the study findings, their potential usefulness for portfolio investors and managers and an indication of questions opened for future research.

REFERENCES
Chapter 1

1. Broad Street Investing Corporation, prospectus, (New York: May 1, 1971), p. 1.
2. Whitehall Fund, Inc., prospectus, (New York: May 1, 1971), p. 1.
3. J. Von Neuman and O. Morgenstern, Theory of Games and Economic Behavior, 3rd edition (Princeton, N.J.: Princeton University Press, 1953), p. 25.
4. P. E. Green and F. J. Carmone, Multidimensional Scaling and Related Techniques in Marketing Analysis, Marketing Science Institute Series (Boston: Allyn and Bacon, Inc., 1970), p. 72, 89.
5. Harry M. Markowitz, "Portfolio Selection," Journal of Finance, Vol. VII (March, 1952), pp. 77-91.
6. Harry M. Markowitz, Portfolio Selection: Efficient Diversification of Investments (New York: John Wiley & Sons, Inc., 1959).
7. William F. Sharpe, "A Linear Programming Algorithm for Mutual Fund Portfolio Selection," Management Science, Vol. XIII (March, 1967), pp. 499-510.
8. Alberto R. Ilano, "Alternative Approaches to Risk in Portfolio Analysis," unpublished Ph. D. dissertation (Berkeley, Calif.: University of California, 1969).
9. G. E. Pinches and W. R. Kinney, Jr., "The Measurement of Volatility of Common Stock Prices," Journal of Finance, Vol. XXVI (March, 1971), pp. 119-125.
10. Keith V. Smith, Portfolio Management (New York: Holt, Rinehart and Winston, Inc., 1971), p. 125.
11. Markowitz, "Portfolio Selection," p. 77 and p. 91.
12. Geoffrey P. E. Clarkson, Portfolio Selection: A Simulation of Trust Investment (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962).
13. Donald E. Farrar, The Investment Decision Under Uncertainty (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962) p. 75 [approximate rendition].

14. L. Fisher and J. Lorie, "Some Studies of Variability of Returns on Investments in Common Stocks," Journal of Business, Vol. XLIII (April, 1970).
15. A. Feeney and D. Hester, "Stock Market Indices: A Principal Components Analysis," in Risk Aversion and Portfolio Choice, ed. by D. Hester and J. Tobin (New York: John Wiley & Sons, Inc., 1967).
16. Benjamin King, "Market and Industry Factors in Stock Price Behavior," Journal of Business, Vol. XXXIX (January, 1966), pp. 139-190.
17. Stephen L. Meyers, "A Factor Analysis Approach to Studying the Structure of the Stock Market," unpublished Ph. D. dissertation (Philadelphia: University of Pennsylvania, 1970).
18. Rudolph J. Rummel, Applied Factor Analysis (Evanston, Ill.: Northwestern University Press, 1970).
19. Farrar, The Investment Decision, p. 35.
20. Roger N. Shepard, "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function - Part I," Psychometrika, Vol. XXVII (June, 1962), pp. 125-140; Part II - Vol. XXVII (September, 1962), pp. 219-246.
21. W. Wagenaar and P. Padmos, "Quantitative Interpretation of Stress in Kruskal's Multidimensional Scaling Technique," British Journal of Mathematical and Statistical Psychology, Vol. XXIV (1971), pp. 101-110.
22. Shepard, "The Analysis of Proximities - Part I," p. 241.
23. Ledyard R. Tucker, "Relationship of Scaling and Factor Analysis," mimeographed paper delivered before the Bell-Penn Workshop on Multidimensional Scaling (Philadelphia: University of Pennsylvania; June 7-10, 1972).
24. D. True and R. Matson, "Cluster Analysis and Multidimensional Scaling of Archeological Sites in Northern Chile," Science, Vol. CLXIX, No. 3951 (1970), p. 1201.
25. P. Green and A. Maheshwari, "Common Stock Perception and Preference: An Application of Multidimensional Scaling," Journal of Business, Vol. XLII (October, 1969), pp. 439-457.

Chapter 2

AN EXPERIMENTAL DESIGN: DATA PREPARATION

The foregoing Introduction has set the stage for a plan of experimentation. Our research objective is to operate on a sample of mutual funds using multidimensional scaling to detect, if possible, factors which mathematically account for differences in their portfolio holdings. This chapter explains the choice of data and sampling rationale.

Before progressing, it would be useful to summarize the procedure of our experiments involving similarity scaling. The nonmetric scaling algorithm takes as input a matrix of proximity measurements between all pairs of the subjects' portfolios; it processes the proximities as rank order distances; and it derives a metric configuration of mutual fund points in a space of prespecified dimensionality such that the interpoint distances best match the input proximities in a monotone regression sense. The output of the computer analysis is a swarm of points alone in space, independent of any particular coordinate system. The configuration of points itself, coupled with external information, should suggest the placement of a meaningful set of reference axes. The appropriate number of axes to be fixed will be indicated by a comparison of the stress and interpretability of alternative solutions over a range of dimensionalities.

Distinct phases of experimental design are suggested in the procedural steps above; namely -

- 1) Selecting the sample of mutual fund subjects
- 2) Measuring the subjects' stimuli response profiles--classifying the funds' portfolio positions into standard investment groups
- 3) Computing pairwise proximities between subjects across their response profiles
- 4) Scaling the matrix of proximities to achieve alternative solutions for a range of prespecified dimensionalities
- 5) Interpreting the solution spaces--determining the appropriate dimensionality and orienting a meaningful coordinate system.

At each of these phases there are many design options available. Following is a description of how the first three phases--the data preparation steps--have been handled for the experimental analysis of mutual fund structure for a single point in time. An explanation of the remaining two phases--the actual single-point-in-time scaling experiment--and the results achieved comprises Chapter 3.

SELECTING MUTUAL FUND SUBJECTS

Selection of a sample of mutual funds must be carried out with an eye toward facilitating the interpretation of scaling results in the final stage of the experiment. That is, a stratified sample has been sought, one which would be representative of the full range of possible mutual fund investment objectives. If trends or clusters are detectable in the

spatial pattern of fund points consistent with known investment objectives, reference axes relating to these objectives could be set.

To lessen the subjectivity that would be involved in trying to classify funds according to their self-stated investment objectives, the decision was made to rely on the classificatory judgments published by professional mutual fund advisory services. Furthermore, to reduce the remaining subjectivity associated with using any single service, the decision was made that a fund would qualify for the sample only if its investment objectives classification was consistent for all of several different advisory services.

Accordingly, four authoritative sources for mutual fund evaluations were used as follows:

- 1) FundScope Magazine [1]
- 2) Arthur Lipper Corporation [2]
- 3) Standard and Poor's Corporation [3]
- 4) Wiesenberger Services, Inc. [4]

The various investment objective classification schemes for these sources were sufficiently alike to allow meaningful comparisons. A consensus system used in the present study consisting of seven categories is described in Table 1. This system comprises an incomplete hierarchical ordering of two primary investment objectives, Income and Growth, and a secondary Stability objective.

Table 1

CONSENSUS MUTUAL FUND CLASSIFICATION
SYSTEM BY INVESTMENT OBJECTIVE CATEGORIES ¹

Symbol	Description
Primary Categories:	
G ₁	"Performance" or "Maximum Capital Gains" Growth Funds
G ₂	"Diversified Stock" Growth Funds
GI	Growth-Income Funds
IG	Income-Growth Funds
I	Income Funds
Secondary Categories:	
GIS	Growth-Income-Stability Funds
SIG	Stability-Income-Growth Funds

1. Mutual funds also are classified as to their investment policies; that is, according to the type of securities in which they invest. A mutual fund that invests more than about 80 percent of its assets in common stocks or similar securities is considered to be a "common stock fund." A fund that invests, say, 20 to 50 percent of its assets in bond-type securities is considered to be a "balanced fund." A "bond fund" typically has a majority of its assets in bond-type securities. Among common stock funds are specialized funds which concentrate their portfolios in securities of a particular industry (e. g., insurance), of a geographic region, a foreign country and so forth. "Hedge funds" seek protection by selling short certain securities at the same time they are long in others.

The sampling population for the present study included common stock, balanced and bond funds and the sample finally selected included funds of each of these policy categories. However, specialized common stock funds and hedge funds were excluded from selection.

As of the beginning of 1971 there existed 586 registered open-end investment companies (mutual funds) of all sizes. Table 2 shows a breakdown of these funds, classified according to investment objectives.

Table 2

Population of Mutual Funds as of December 31, 1970¹

<u>Objective of Fund</u>	<u>Number of Funds</u>	<u>% of Total Number</u>
G ₁	189	32.3
G ₂	168	28.7
G ₁ and IG	118	20.1
I	36	6.1
SIG and GIS	30	5.1
Specialized and others not classified	<u>45</u>	<u>7.7</u>
	586	100.0

The maximum number of funds for the total sample in the present study was constrained by computational considerations. Computer memory and processing time limitations as well as specifications of the scaling programs that were used effectively limited the number of subjects to 46. A sample of this size was more than enough to assure virtually unique solutions for the range of dimensionalities studied (up to five).

1. Adapted from Wiesenberger's classification system [5].

Another consideration in selecting a sample of subjects was to have a representation of funds of different sizes ranging up to the largest having over \$2 billion in total assets.

Within these guidelines the sample of mutual funds was chosen randomly from the universe of all funds with assets over \$25 million. The sampling of 46 subjects was allocated to 8 each for the 5 primary investment objective strata and 3 each for the 2 secondary Stability-related categories. Table A1 names the funds selected. These 46 funds represent 17 percent of the total number of investment companies (open- and closed-end) larger than \$25 million in existence as of March 31, 1971, and 29 percent of all investment companies' total assets.

PROFILE DATA

Forty-six mutual funds represent the subjects for the scaling experiment. The universe of possible fund holdings represents the experiment's test stimuli. Therefore, a fund's portfolio of actual holdings can be considered the subject's profile of stimuli responses.

In order to reduce the stimuli universe of possible fund investments to manageable proportions, it is necessary to aggregate portfolio holdings into standard industry and asset groupings. However, a drawback is apparent. Two mutual funds, each which uniquely holds one security in the same industry group but of different companies, say Standard Oil of New Jersey and Occidental Petroleum Company, would

be said to be completely similar despite the probable judgment that their component investments have important perceived differences. In spite of this, aggregation into industry and asset groups makes sense for several compelling reasons.

First, a comparison on the basis of individual companies would be prohibitive from a computational standpoint. The universe of companies whose securities are eligible to be held by mutual funds is several thousand.¹ Computing pairwise similarities between funds would involve a large number of comparisons for every pair. More important, however, is the fact that a huge number of different funds would have to be compared to produce meaningful scaling results--undoubtedly more funds than actually are extant.

Second and most important, comparisons on the basis of individual companies would not capture the industry effect in similarity perceptions. Thus, two mutual funds which uniquely hold securities of General Motors Corporation and Ford Motor Company, respectively, would be judged to be completely dissimilar.

The existence of a strong industry effect in investment evaluations is a feature of investor behavior that has been reported many times. As Clarkson has observed, investors associate specific industries

1. Arthur Lipper Corporation estimates that the number of different common stocks held by the mutual fund industry at the beginning of 1971 was approximately 2700 [6].

with different investment objectives. Smith has expressed the idea that industry membership is a major consideration in selecting securities:

"The general theme ... to diversify across the favorable industrial categories by investing equal amounts in popular securities within each category ... is a prescription for portfolio selection which turns out to be an apt description of the intuitive decision-making process prevailing in practice." [7]

Also, in commenting on the results of a study by Friend and others [8],

Cohen and Zinbarg have stated:

"Investment company managers appear to be willing to make major decisions on industries in which to concentrate but they also appear to seek the safety of numbers by purchasing many individual issues within the industries." [9]

In short, the aggregation of mutual fund holdings into industry and asset groups is not only the sole practicable solution but is also justified by the nature of investor behavior. Organizing mutual fund holdings into such groups has been reported as a first step in empirical studies by Friend [10] and Smith [11] as well as Farrar's work cited earlier (page 20).

For these reasons, a set of standard investment groups was defined consisting of 94 industry common stock categories plus one category each for fund holdings in preferred stocks, long-term debt securities and cash. Table A2 lists the 97 groups. A criterion for defining the groups was that they should reflect industry distinctions as recognized by security analysts and investors generally.

Several ready-made industry classification systems for investment purposes are available. For example, Standard & Poor's Corporation publishes statistics for their scheme of 100 industry groups.¹ The scheme chosen for the present study is based upon a well-suited system provided by the Arthur Lipper Corporation. Since 1971 the Lipper organization has published quarterly compilations of mutual fund holdings called INSIGHT [13] which, among other things, includes a breakdown of each fund's common stock holdings into 103 industry groups. These were designed to reflect the industry groups in which mutual funds typically present their holdings in their prospectuses and reports to shareholders.

Each of Lipper's groups was set up to correspond to one or more 3-digit SIC industries.² Where it has been done, the basis for aggregating SIC industries was to reflect financial homogeneity as distinguished from the SIC purpose of delimiting trade or economic homogeneity. Thus, for example, gold mining companies (SIC-104) were treated as a category separate from other metal mining companies (SIC-101 to 109) in accordance with the peculiar financial features of gold mining securities.

To comply with constraints on the number of variables imposed by the scaling computer programs that were used, the original 103

1. See Standard & Poor's Corporation's "Analysts Handbook" [12].

2. Re: Standard Industrial Classification code published by the United States Bureau of the Budget [14].

industry groups were reduced to 94. Nine groups, which were subsequently shown to be under-represented in the portfolios of the chosen sample of 46 mutual funds, were eliminated by combining them with other groups which were judged to be similar from an investment viewpoint.¹

Any comparative analysis of the holdings of a sample of mutual funds would be meaningful only if their portfolios represent the same point in time. March 31, 1971, the end of the first quarter reporting period in that year for mutual funds to the Securities and Exchange Commission, was the date chosen. For each of the 46 subject funds at this date, the percentage of total fund assets represented by each industry and asset group holdings was obtained from the Arthur Lipper INSIGHT publication.²

1. At the beginning of 1971 less than two percent of the total assets of all investment companies over \$25 million in size represented securities held in these nine eliminated categories. Four of these categories were not represented at all in the holdings of the 46 fund test sample.

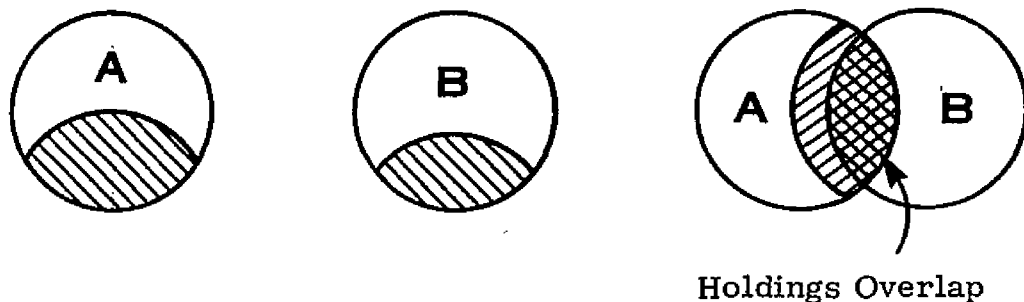
2. Arthur Lipper's classification of common stock holdings into industry groups included only securities held of companies listed on the New York and American Stock Exchanges. Although these securities represented a large part of the total value of common stocks held--approximately 88 percent for the fund sample--it was deemed desirable to revise the percentage of total asset values to include all securities traded over-the-counter. The reason was that certain industries, particularly insurance and banking, have a disproportionately high representation in over-the-counter securities. Accordingly, for each fund the percentage of total assets represented by each over-the-counter issue was added to its respective industry group. The appropriate industry group was determined from the SIC identification for the issuing company according to Security and Exchange Commission classifications [15].

The average mutual fund in the sample held common stock securities in 92 individual companies representing 33 industry groups. Table A3 presents the resulting data matrix of 46 funds by 97 investment groups. Altogether, this tabulation of First Quarter 1971 holdings represents the classification of approximately 4300 separate investment positions.

A PROXIMITY MEASURE

The final step in preparing the basic data for similarity scaling was to compute the pairwise associations between funds in terms of the communalities of their respective portfolio profiles in Table A3. A special-purpose measure of association was designed called "holdings overlap."

The holdings overlap between any two funds is a measure of the maximum percentage of dollar value they both hold in the same investments. For an artificial example, assume portfolios A and B only have one industry or asset group in common. If A's holding in that group represents 35 percent of A's total assets and B's holding in that group represents 20 percent of B's total assets, then A and B's overlap is 20 percent, the minimum of 35 and 20 percent. This can be represented in a Venn diagram:



Holdings overlap is the intersection of two sets ($A \cap B$). It is computed by the following formula:

$$S_{ab} = \sum \min (h_{ai}, h_{bi})$$

where S_{ab} stands for the similarity between funds A and B and h_{ai} and h_{bi} are the funds' respective percentage of total portfolio assets invested in investment group i, taken from Table A3. Similar overlap or matching coefficients have been used before as in the study of verbal stimuli [16], and in this context they have produced "very good results" [17].

Holdings overlap computed for all pairs of mutual fund portfolios comprises a symmetric matrix of values ranging from a possible 0.0 to 1.0. Along the matrix diagonal are 1.0's for a portfolio's overlap with itself. This means that holdings overlap properly is to be considered a measure of similarity (standing not for distances but complements of distance). Table A4 presents the lower half-matrix (without diagonal) of 1058 overlap values computed for the 46 mutual fund subjects in the experiment. This table represents the basic input for the multidimensional scaling algorithms.

REFERENCES
Chapter 2

1. "Mutual Fund Group Classifications," FundScope, (April, 1971), pp. 200-201.
2. Arthur Lipper Corporation, "Mutual Fund Performance Analysis," a weekly bulletin (New York, March 31, 1971).
3. Standard & Poor's Corporation, "Mutual Fund Summary," Stock Guide (New York, March, 1971), pp. 235-249.
4. Wiesenberger Services, Inc., Investment Companies 1971 (New York, 1971), pp. 99-111.
5. Ibid., p. 44.
6. Arthur Lipper Corporation, "Holdings and Transactions; First Quarter, 1971," Book One of "INSIGHT" (four books; New York, 1971), p. v.
7. Keith V. Smith, Portfolio Management (New York: Holt, Rinehart and Winston, Inc., 1971), pp. 125-126.
8. Irwin Friend et al, "A Study of Mutual Funds," prepared for the S. E. C. Report of the House Committee of Interstate Commerce & Finance, 87th Congress, 2nd Session (U. S. Government Printing Office, 1962).
9. J. Cohen and E. Zinbarg, Investment Analysis and Portfolio Management (Homewood, Ill.: Richard D. Irwin, Inc., 1967), p. 625.
10. I. Friend and D. Vickers, "Portfolio Selection and Investment Performance," Journal of Finance, Vol. XX (September, 1965), pp. 391-415.
11. K. V. Smith and J. C. Schreiner, "A Portfolio Analysis of Conglomerate Diversification," Journal of Finance, Vol. XXIV (June, 1969), pp. 413-427.
12. Standard & Poor's Corporation, "Analysts Handbook" (New York, published annually).
13. Arthur Lipper Corporation, "INSIGHT," Book Four.

14. United States Bureau of the Budget, "Standard Industrial Classification Manual," (U. S. Government Printing Office, 1971).

15. United States Securities and Exchange Commission, "Directory of Companies Filing Annual Reports with the Securities and Exchange Commission," (U. S. Government Printing Office; June, 1969).

16. James Deese, "On the Structure of Associative Meaning," Psychological Review, Vol. LXIX (May, 1962), pp. 161-175.

17. J. Douglas Carroll, private communication, May, 1970.

Chapter 3

THE EXPERIMENT

The core experiment of this study is the multidimensional scaling of the proximities matrix of the last chapter representing derived similarities between a sample of mutual funds at a single, given point in time. As stated in the Introduction, many options exist as to the manner of scaling including the choice of algorithm; whether the input data are treated metrically or nonmetrically; whether a Euclidean or non-Euclidean solution space is specified, and so forth. Chapter 5 examines the sensitivity of scaling results to variations in the major options for data preparation and scaling experimentation. This chapter presents the experimental design which, as a consequence of deduction and trial and error search, was found to give clearly the most successful results.

The matrix of proximity data in Table A4 captures the relationships between the portfolios of mutual fund subjects. Every possible set of three data values can be used with the triangle inequality axiom of metric space as an inequality constraint on the distances between subject points in a spatial scaling solution. Thus, starting with a randomly generated solution configuration in a space of prespecified dimensionality, subject points can be moved about in such a manner as to reduce the cardinal extent of violations of the ordinal inequality

relationships. "Stress" of a solution configuration is a measure of the degree to which the configuration violates the rank order constraints of the original input data. Procedures can be designed to move systematically the subject points to converge on a solution of least possible stress. With as few subjects as three times the number of dimensions of the solution space, enough inequality constraints exist to restrict the movement of subject points so that, for all practical purposes, a unique configuration is obtained.

This process for reaching a solution configuration is embodied as an iterative procedure in the highly versatile computer program M-D-SCAL. M-D-SCAL version 5M was used for all similarity scaling experiments in the present study.¹

M-D-SCAL analyses were run using as input the matrix of holdings overlap proximities between the 46 mutual fund subjects. Alternative solutions were called for in Euclidean spaces from one to five dimensions. Actually, the computer program starts with the space of highest dimensionality, finds a best-fit solution, then uses the projections of this solution as a starting configuration for an analysis of the space of next lowest dimensionality.

1. For a description of the theory and features of the M-D-SCAL computer program, see Kruskal [1] and Green and Rao [2]. Other programs available for the same purpose of similarity multidimensional scaling include TORSCA [3] and KYST [4]. KYST is a recently announced program which represents a refinement of M-D-SCAL V and TORSCA 9.

For each of the solution spaces requested, a configuration of mutual fund points was produced, the fund point positions expressed as coordinates in an arbitrary reference system. Table A5 presents the coordinate results for the alternative solutions in one-, two-, and three-space. A graphical display of the two dimensional solution, as it is produced by the M-D-SCAL program, is shown in Figure 2.¹ Each fund is labeled as to its investment objective category (from Table A1).

The intent of calling for multiple solutions of ascending dimensionality was to detect, if possible, the most appropriate solution space dimensionality. As dimensions are added, the stress should drop noticeably until the number of factors are reached which account for major sources of variations in the data. Thereafter, stress should level off to a slow decline with increasing dimensionality. Ideally, a plot of stress versus dimensionality would show a sharp elbow at the dimensionality of the "correct" solution space (see the ideal stress curve in Figure 3, page 67).

Although in multidimensional scaling studies of subjects and stimuli that correspond to physical variables a distinct crimp often can be detected in the stress diagram at the appropriate dimensionality, in most studies of human subjects and behavioral stimuli the stress

1. The graph as it was produced by the M-D-SCAL program had the upward direction in Figure 2 pointed downward. The original display has been flipped over (reflected) in Figure 2 for ease of comprehensibility. This is an allowable transformation, equivalent to multiplying the Y coordinate values in Table A5 by a proportionality constant (-1).

65

M-D-
BASED

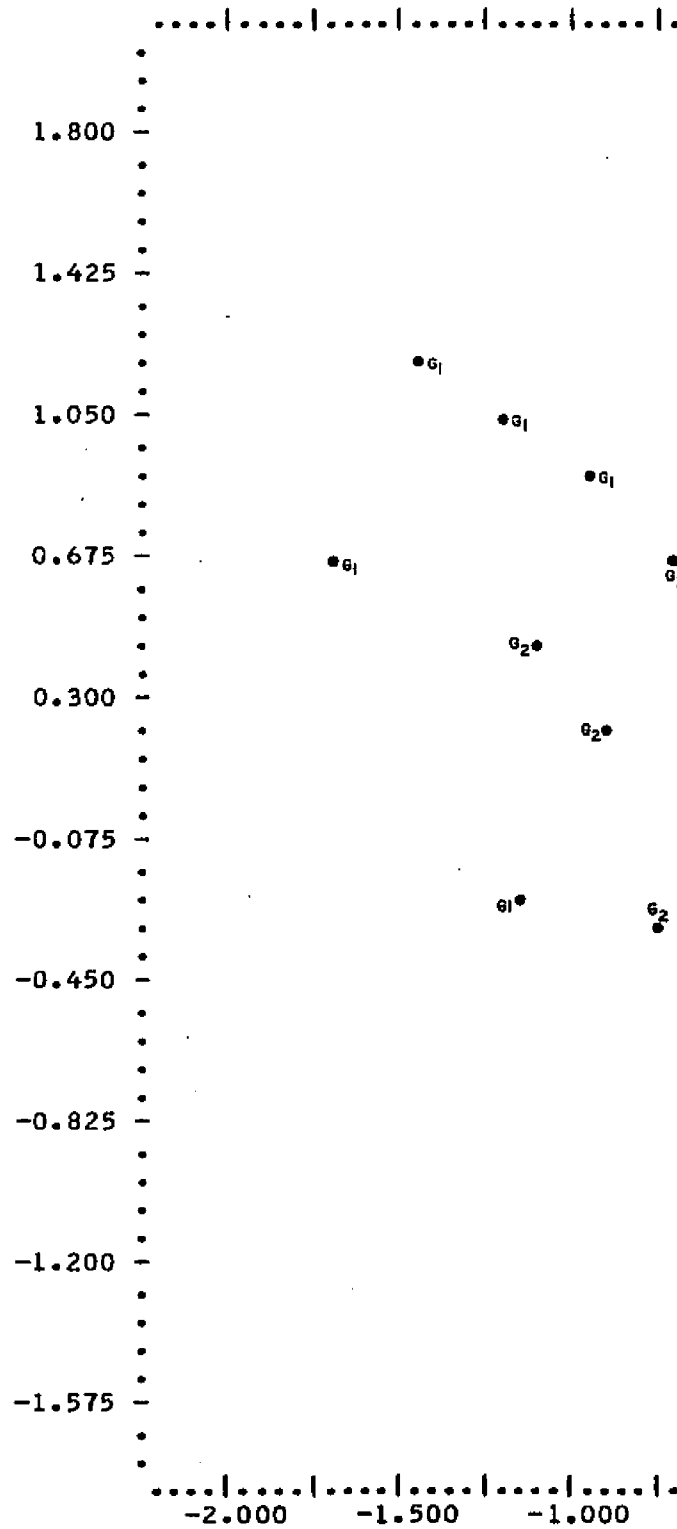


diagram forms a relatively smooth convex trend which does not provide clear indication of the choice of dimensionality. In the absence of a dominant corner in the stress diagram, the selection of a solution space becomes a balance between goodness-of-fit and simplicity of representation. Despite perhaps a good improvement in stress, the addition of an extra dimension would be undesirable if it impairs the comprehensibility of a solution. Goodness-of-fit often has been compromised to achieve parsimony and interpretability in a multidimensional scaling display of data.

Stress values obtained in the scaling experiments for one through five dimensions are tabulated below in Table 3 and are shown graphically in Figure 3.

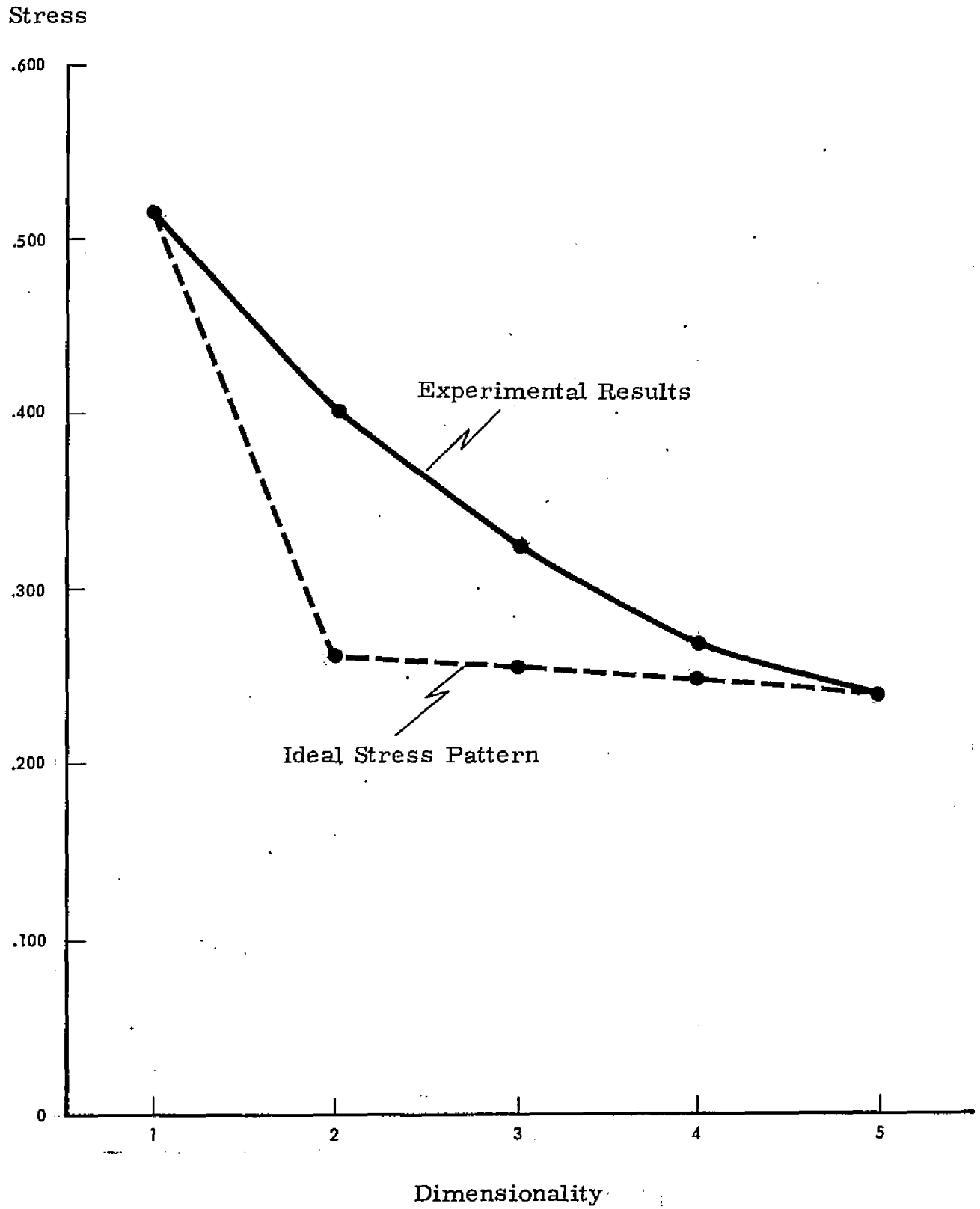
Table 3

<u>Dimensionality</u>	<u>Stress</u>
1	0.514
2	0.400
3	0.325
4	0.269
5	0.242

The question, "What constitutes an acceptable level of stress," remains an unsettled issue in multidimensional scaling theory. It is affected by many conditions of the particular experiment. Joseph Kruskal has offered some guidelines for appraising the quality solution results when using M-D-SCAL under the conditions of the present study;

Figure 3

M-D-SCAL SOLUTION STRESS
FOR ALTERNATIVE DIMENSIONALITIES



namely, for nonmetric scaling of a half-matrix (with diagonal absent) having only a single replication. For a range of dimensionality from two to five, Kruskal's appraisals are as follows [1]:

Table 4

Poor	0.40
Fair	0.20
Good	0.10
Excellent	0.05
"Perfect"	0.00

This table applies to sample sizes from about ten to thirty subjects. Higher stress is more acceptable for larger samples as in our study of 46 subjects. Nonetheless, the values we obtained are quite high though this does not necessarily lead to the conclusion that we should focus on the higher dimensionality, lower stress solution spaces. As explained in the Introduction, the level of stress for any solution space is a combination of two effects: 1) lack of representation of all dimensions, and 2) noise in the proximity data. If we look at a stress-dimensionality curve, the level of the curve is related to the error effect and the shape of the curve is related to the dimensionality effect. Because our stress curve indicates slow convergence at high stress values, we can conclude that an important contribution to the stress readings is noise in our input data. Chapter 5 examines the implications of our obtained stress values and discusses alternative scaling experiments giving solutions of substantially less stress.

Considering the shape of our obtained stress curve, the increase from one to two factors shows clear drop in stress. Although the addition of a third variable may contribute to our solution, beyond three dimensions the improvement is doubtful. Accordingly, we shall explore the meaning of the two dimensional solution and then see what, if any, information is added by going to higher level solutions.

Solution Interpretation - Cluster Analysis

If an interpretation for our solution can be discerned, this would lend credence to the experimental methodology, especially if patterns which have been anticipated emerge in the solution space. An analytic technique which is inappropriate for a problem would not be expected to produce solutions having recognizable characteristics. We should check the solution configurations to see if, as expected, the hierarchy of investment objective categories emerges in the spatial pattern.

Even a cursory examination of the two dimensional scaling solution in Figure 2 reveals the tendency for mutual funds belonging to the same class of investment objectives to occur in the same region in the solution space. The extent of this apparent clustering can be brought out by drawing contours to enclose convex sets of mutual funds which are homogeneous as to their investment objectives. The clusters can be drawn in a hierarchical fashion. Figure 4 shows two convex clusters which, at the most fundamental level in the hierarchy of distinctions, correspond to Growth-oriented funds (G_1 , G_2 , G_I , G_{IS}) versus Income-

Figure 4

M-D-SCAL SOLUTION SPACE--MARCH 31, 1971
SHOWING CLUSTERING OF GROWTH AND INCOME
ORIENTED MUTUAL FUNDS

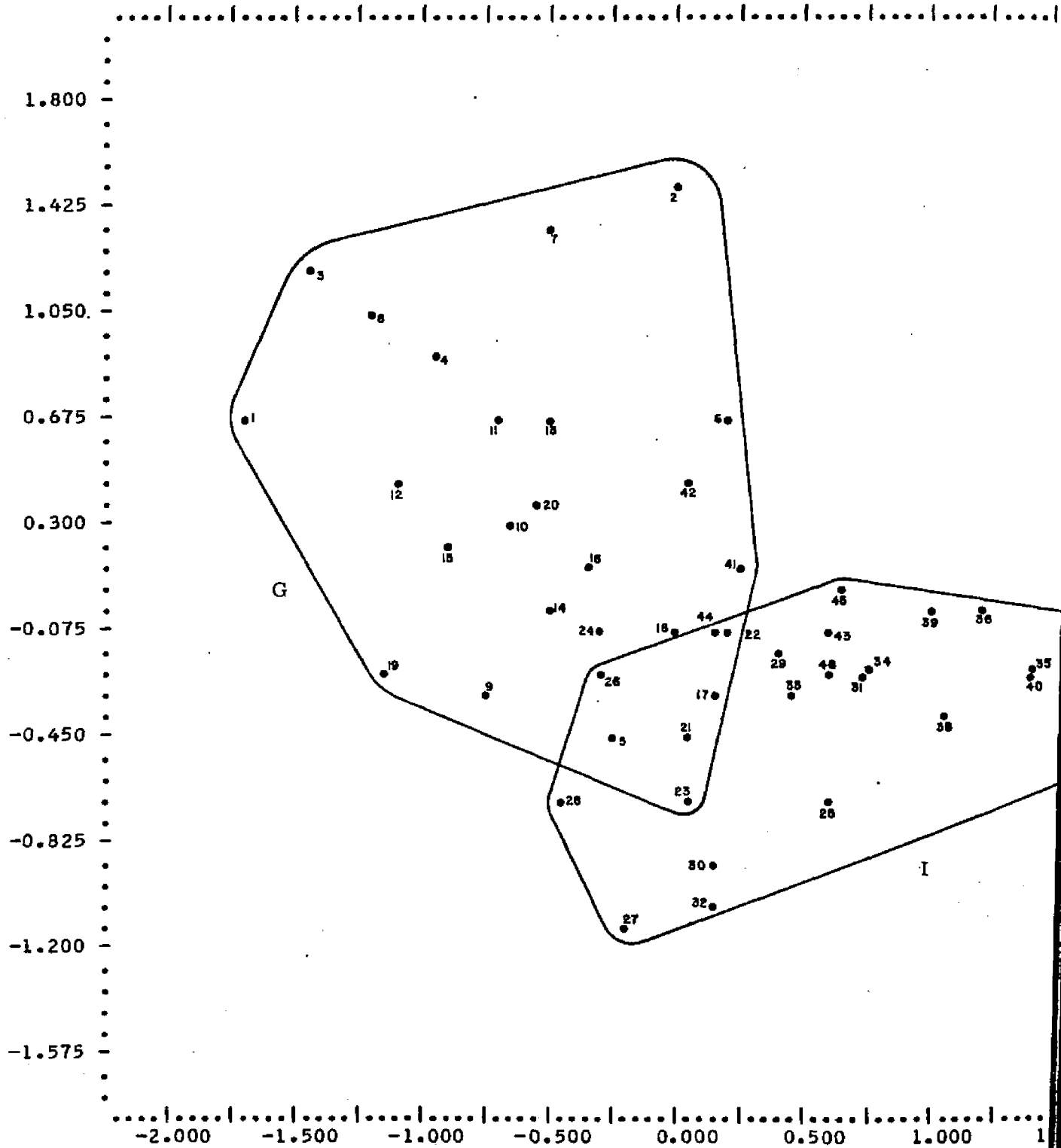
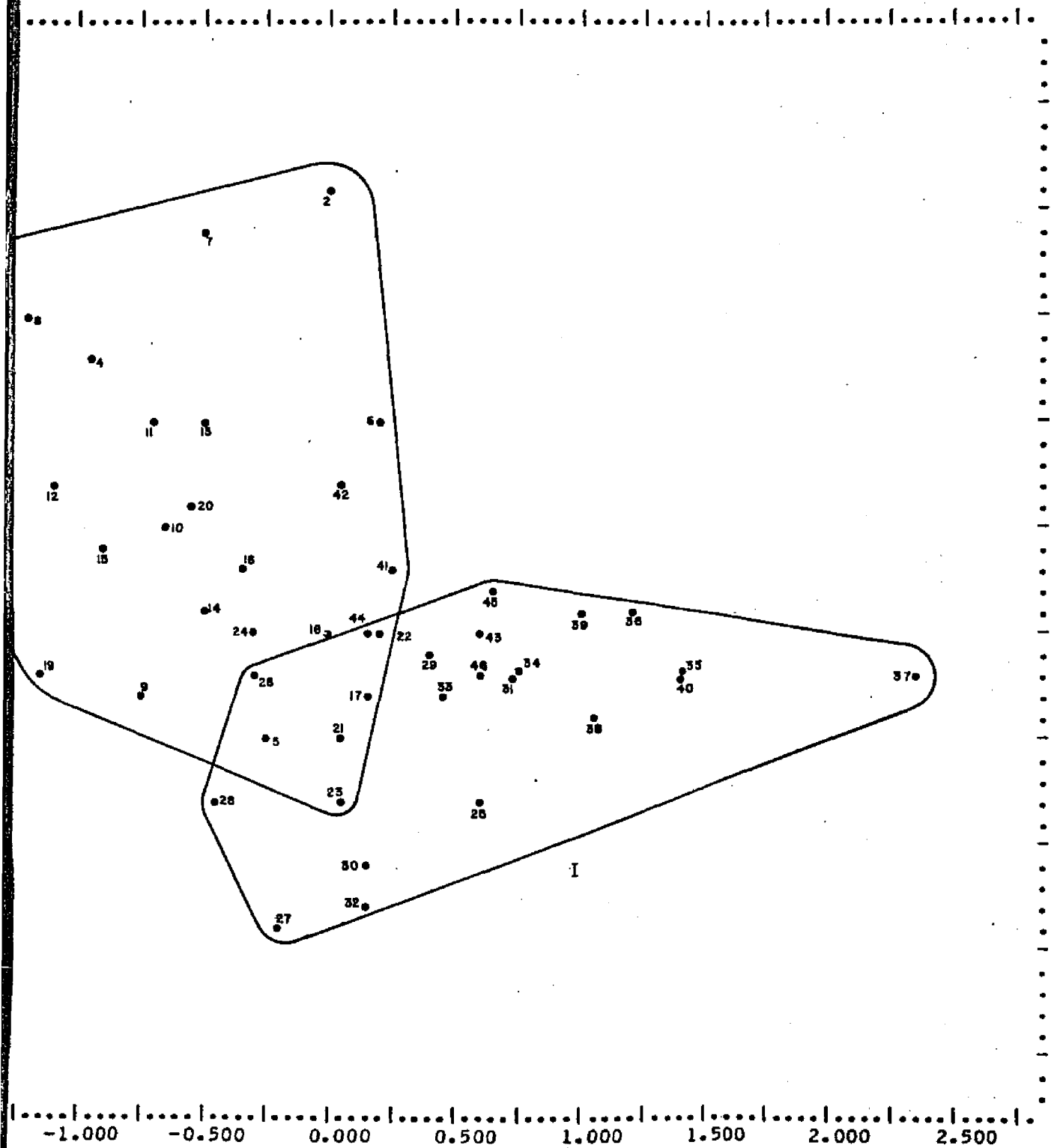


Figure 4

M-D-SCAL SOLUTION SPACE--MARCH 31, 1971
SHOWING CLUSTERING OF GROWTH AND INCOME
ORIENTED MUTUAL FUNDS



oriented funds (I, IG, SIG). Another graph has been drawn, Figure 5, to show clustering at the second level in the hierarchy of investment objective categories for mutual funds.

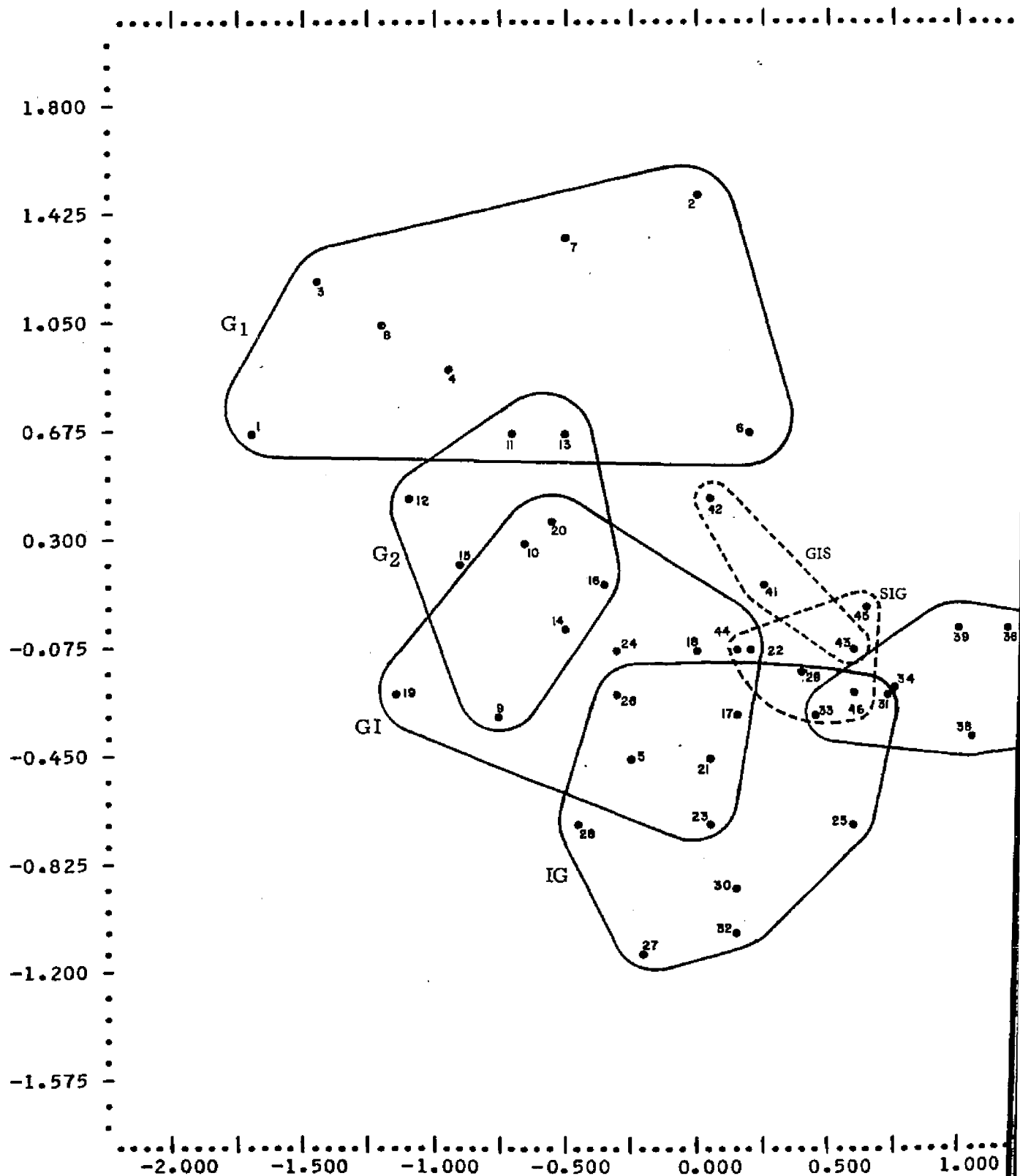
In Figure 5 it can be seen that clusters of mutual funds, homogeneous as to investment objectives, emerge in a remarkably well-defined manner. As would be expected, the G_1 and I clusters fall at extreme ends of the pattern apart from each other with the combination categories arrayed in correct, though overlapping, order in between.

The appearance of a hierarchy of clusters in agreement with our prior expectations of what the solution space should contain tends to support the validity of the methodological approach. However, it is fair to ask whether our focus on pre-supposed results might not have diverted attention from other, more dominant but unanticipated patterns in the solution space. In other words, clusters may be definable in the solution space which are more natural and which, perhaps, relate to some unknown factors that better account for variations in the data. To answer this calls for an objective cluster analytic technique, one that is unencumbered by prior assumptions as to the factors responsible for clustering.

Johnson's hierarchical clustering algorithm, HICLUST [5], is just such a suitable technique. Its built-in objective is not just to partition a heterogeneous set of points in some space into local homogeneous clusters, but also to give a hierarchy of groups of clusters

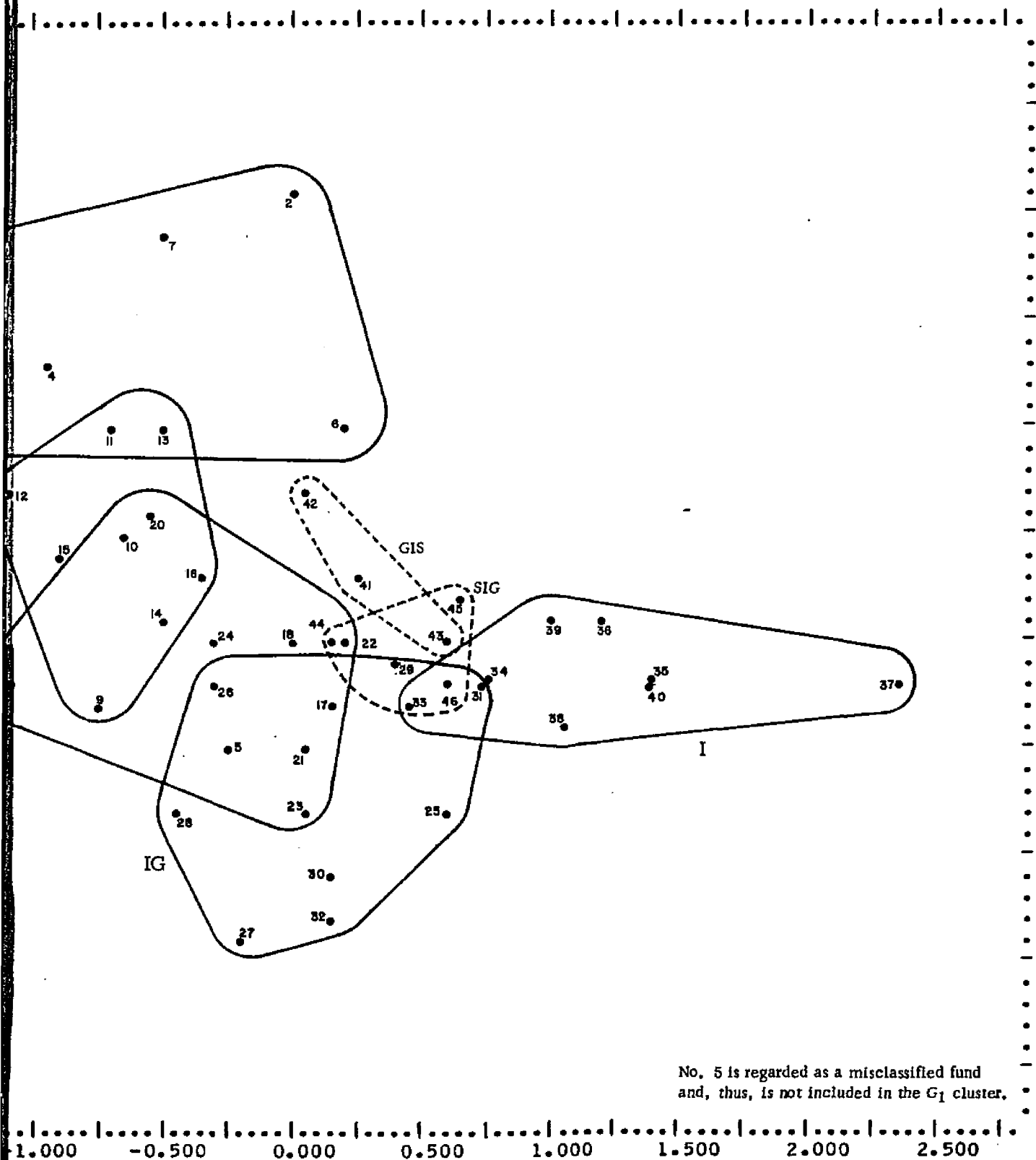
Figure 5

M-D-SCAL SOLUTION SPACE--MARCH 31, 19
SHOWING MUTUAL FUNDS CLUSTERED IN ACCOR
WITH THEIR INVESTMENT OBJECTIVE CLASSIFIC



-72-
Figure 5

M-D-SCAL SOLUTION SPACE--MARCH 31, 1971
SHOWING MUTUAL FUNDS CLUSTERED IN ACCORDANCE
WITH THEIR INVESTMENT OBJECTIVE CLASSIFICATIONS



until at the end all clusters are merged into one universal group. Input to the HICLUST computer program are distance measures or similarities between points and are treated nonmetrically. This means that the clustering solutions are invariant under a monotonic transformation or, otherwise stated, are dimension-free.

Euclidean distances were computed between all the subject points in the M-D-SCAL two-space solution of Figure 2. This triangular matrix of interpoint distances was subjected to a HICLUST analysis (using the program's "diameter" option). Figure A6(a) presents the dendritic HICLUST output showing how the higher level local clusters tend to merge into larger, more basic groups.

HICLUST results have been drawn as contours in the two-dimensional M-D-SCAL solution space. Figure 6 shows the lowest level in the hierarchy of "natural" cluster separations (two groups) and should be compared with the a priori or pre-supposed clustering in Figure 4. Figure 7 shows the four group hierarchical level of natural clusters, the grouping most similar to the a priori clustering in Figure 5.

Despite the fact that the HICLUST procedure does not allow clusters to intersect, a clear correspondence is seen between natural clusters in the scaling solution and the pre-supposed clusters representing investment objective categories. Clusters from the HICLUST analysis can be labeled in a best-judgment, maximum likelihood

Figure 6

NATURAL CLUSTERS AT THE TWO GROUP LEVEL
IN THE TWO DIMENSIONAL M-D-SCAL
MUTUAL FUND SOLUTION SPACE

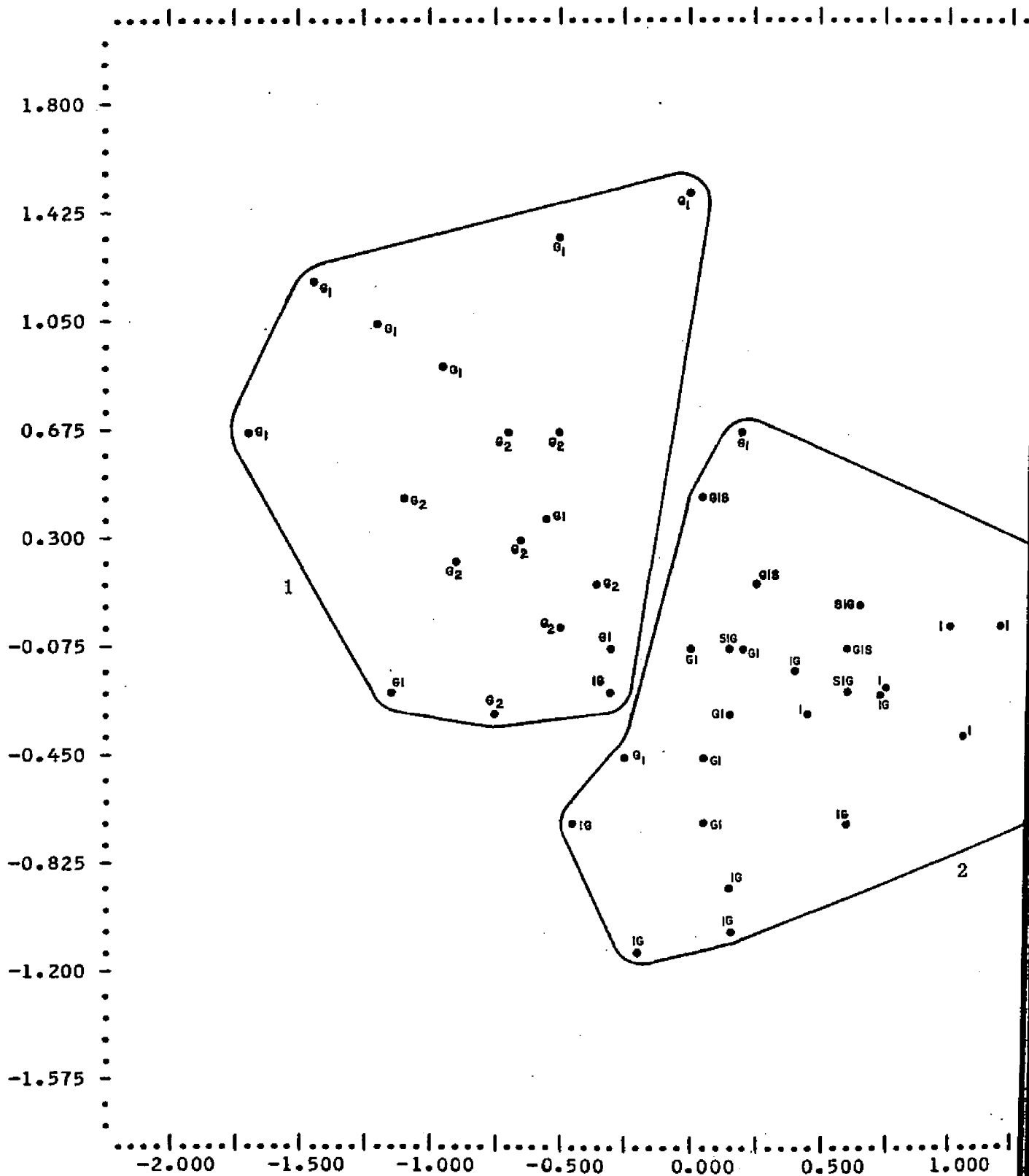
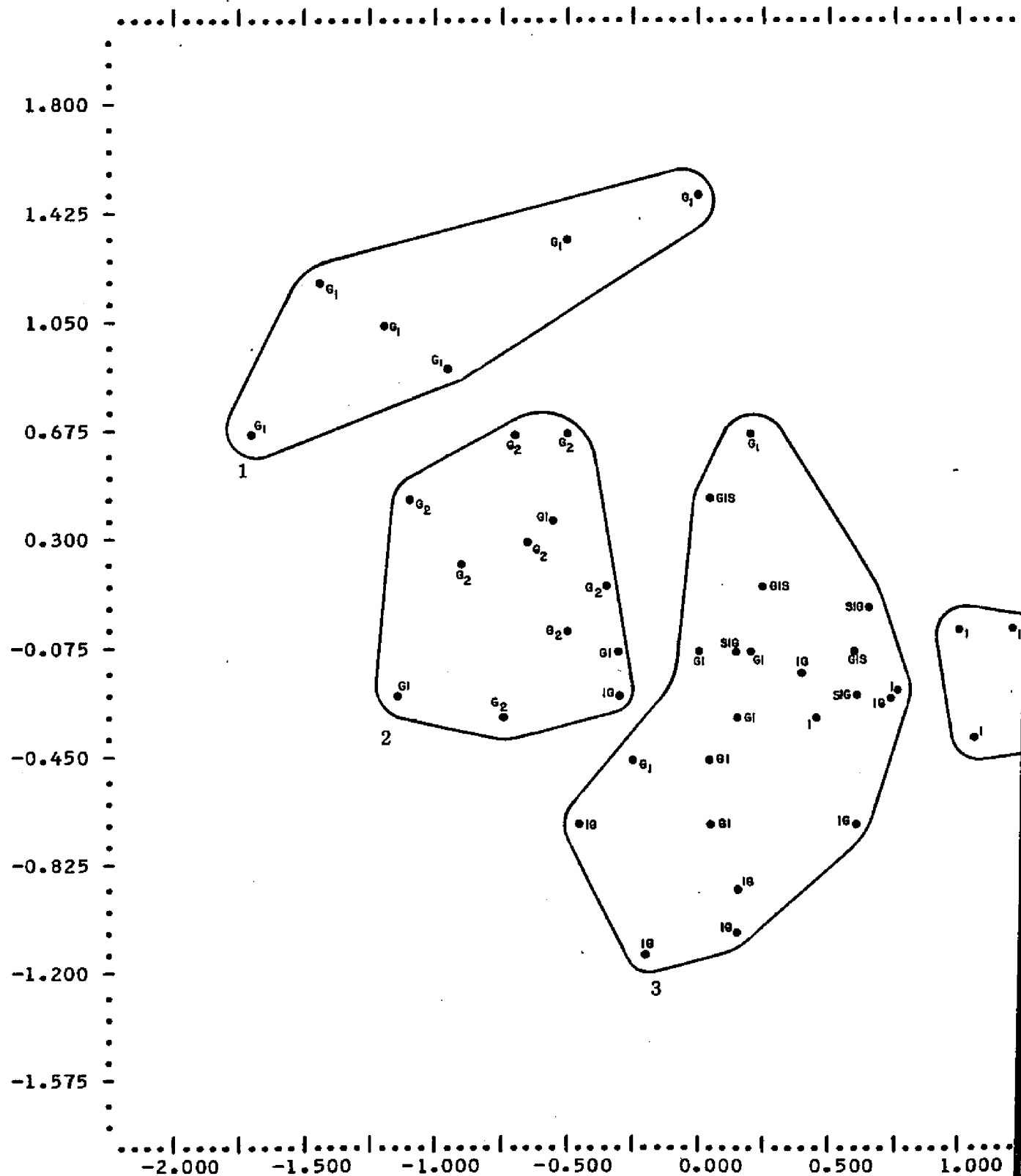


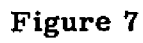
Figure 6

Figure 1 is a scatter plot showing the distribution of 100 points in a 2D space. The points are labeled with 'G1', 'G2', 'IG', 'SIG', and 'I'. The plot is divided into three regions by two curved lines. The top-left region contains points labeled G1 and G2. The bottom-left region contains points labeled IG and G1. The large central region contains points labeled IG, SIG, and I. The x-axis ranges from -0.500 to 2.500, and the y-axis ranges from 0 to 1.000.

Figure 7

NATURAL CLUSTERS AT THE FOUR GROUP LEVEL
IN THE TWO DIMENSIONAL M-D-SCAL
MUTUAL FUND SOLUTION SPACE





sense as to the investment objective category which they correspond. For example, cluster 3 in Figure 7 can be identified as the IG cluster because more of its member funds are of the Income-Growth type than any other variety. Clearly, clusters 1 and 4 can be labeled as the G_1 and I groups because every one of their enclosed funds happen to be of the G_1 and I varieties, respectively. G_2 and GI funds seem to merge into the single cluster number 2 which can be labeled GI. Taken together the four natural clusters in Figure 7 tend to reflect the simple hierarchical scheme of G, GI, IG and I fund categories.

Goodness-of-correspondence between the a priori clustering of Figure 5 and the natural clustering of Figure 7 can be cast into a useful index for comparing alternative scaling solutions. For any natural cluster which has been given an investment objective label, the percentage of member funds in that cluster which, indeed, are of that category type is the goodness-of-correspondence index. In other words, the index is the maximum likelihood probability of the given label for the natural cluster. Table 5 gives the goodness-of-correspondence indices for results shown in Figures 6 and 7.

Table 5

DEGREE OF CORRESPONDENCE BETWEEN
NATURAL AND PREDICTED CLUSTERING
FOR TWO DIMENSIONAL M-D-SCAL SOLUTION

Natural Cluster	Maximum Likelihood Investment Objectives	Maximum Likelihood Probability
<u>Primary Hierarchical Level (See Figures 6 and A6(a), row 2)</u>		
Cluster 1	G (inc. GIS)	0.944
Cluster 2	I (inc. SIG)	0.643
Overall		0.761
<u>Secondary Hierarchical Level (See Figures 7 and A6(a), row 4)</u>		
Cluster 1	G (G ₁)	1.000
Cluster 2	GI (inc. G ₂ & GIS)	0.917
Cluster 3	IG (inc. SIG)	0.455
Cluster 4	I	1.000
Overall		0.696

The test for natural clusters in the solution space confirms the pre-supposed clustering to a gratifying degree. Because patterns have emerged in our solution as expected, we can reasonably conclude the following:

- 1) The experimental design is meaningful.
- 2) Latent factors accounting for variations in mutual fund portfolios relate to fund investment objectives.
- 3) The dominant descriptors of investment objective clusters at the two dimensional level are Growth and Income.

The a priori clusters in Figure 5 do not fall into a straight line as they would if Growth and Income represent the opposite poles of a unidimensional factor which was the source of variations in the input proximities data. Instead, the arc-like pattern of clusters compels us to conclude that at least a two-space representation is necessary, where Income is one factor generally corresponding to the horizontal trend in Figure 5 and Growth is a separate factor approximately corresponding to the vertical. This is brought out in Figure 8 as a plot of the centroids of the a priori clusters from Figure 5.

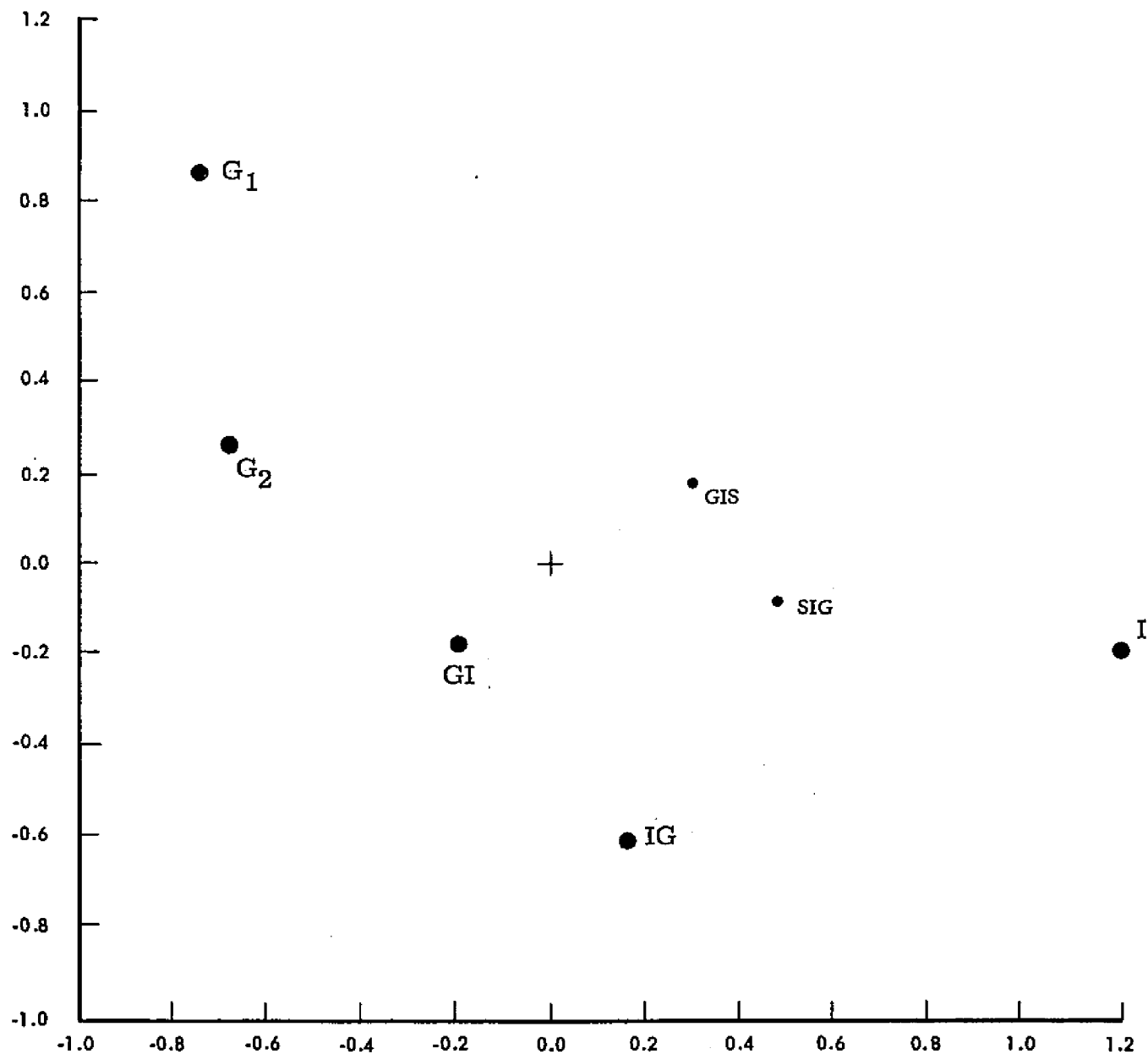
As previously mentioned, the solution space of Figure 5 is as it comes from the M-D-SCAL V computer program, i.e., having an arbitrary orientation. The reference axes are generated by the computer program for convenience of display with the (0,0) point located at the overall configuration centroid. Rulings on the coordinates are interval scaled, implying that only differences between values have meaning.¹

Although the orientation of the reference axes identified with Income (abscissa) and with Growth (ordinate) seems to explain cluster patterns in Figure 8 quite well, some discrepancies appear. The I cluster seems to rate too high on the Growth attribute especially in comparison to cluster IG. Also, the Growth ratings of the GIS and SIG

1. Interval scales, such as common temperature scales, are unique up to a linear transformation, i.e., of the form $y = a + bx$. An example is the conversion of Fahrenheit to centigrade units.

Figure 8

M-D-SCAL SOLUTION SPACE
SHOWING CENTROIDS OF MUTUAL FUND CLUSTERS
(See Figure 5)



clusters seem too high compared to the GI cluster. These discrepancies call for a fine-tuning of the coordinate system through orthogonal rotation so that scale values of the projections on the axes are maximally meaningful in a descriptive and interpretive sense. However, at this point we have about depleted the informational content of the matrix of proximities to find and interpret a factor structure. To get a better fix on the placement of a system of reference axes, we must bring to bear external information concerning known attributes of our mutual fund subjects. An approach to doing this involves the use of what are known as property vectors.

Solution Interpretation - Property Vectors

If the true factors of portfolio discrimination were known and these factors corresponded to mensurable portfolio attributes, we could plot the portfolio of each mutual fund as a point in a space spanned by the known factors. Projections of a point on the axes would correspond to the mutual fund's respective attribute ratings. Thus, as we moved along the Income axis, we would find the point projections of portfolio yields increasing monotonically. Also, moving orthogonally in the direction of Growth axis, we would find the projection values corresponding to expected growth increasing monotonically.

The foregoing suggests a procedure for fitting a coordinate system when all we have is a swarm of points. For each of our subject mutual funds, we obtain a scale value on some property which we believe

corresponds to a basic factor of evaluation. Then it is possible to draw a vector in the solution space in a direction such that projections of the subject points onto this vector correspond maximally with the original property ratings. Different fund properties or attributes can be used as candidate vectors to find factors which best explain the solution space. This is equivalent to a multiple regression with each property scale serving, in turn, as a dependent variable. The notion of "maximum correspondence" alternatively may be defined in terms of linear, non-linear or monotonic regression.

With the property vector approach in mind, an attempt was made to obtain property values corresponding to each subject mutual fund's income and growth attributes. The data source was Mutual Fund Performance Analysis [6], a service published weekly by the Arthur Lipper Corporation. It lists for all registered mutual funds their respective current yields and their percentage change in per-share net asset value over periods of various duration.

Measuring a fund's growth attribute proved problematic. In accordance with both normative and positive models of portfolio selection, property ratings should relate to a portfolio's expected growth. However, because of the great difficulty in measuring growth expectations, ex post growth properties were taken as surrogates for ex ante measurements. Candidate growth properties tested were, alternatively, the percentage change in a fund's per-share asset values that was

actually realized over the three, six, nine, and twelve months following March 31, 1971. Table A7 presents the growth percentages for the 46 mutual funds as well as their yields.

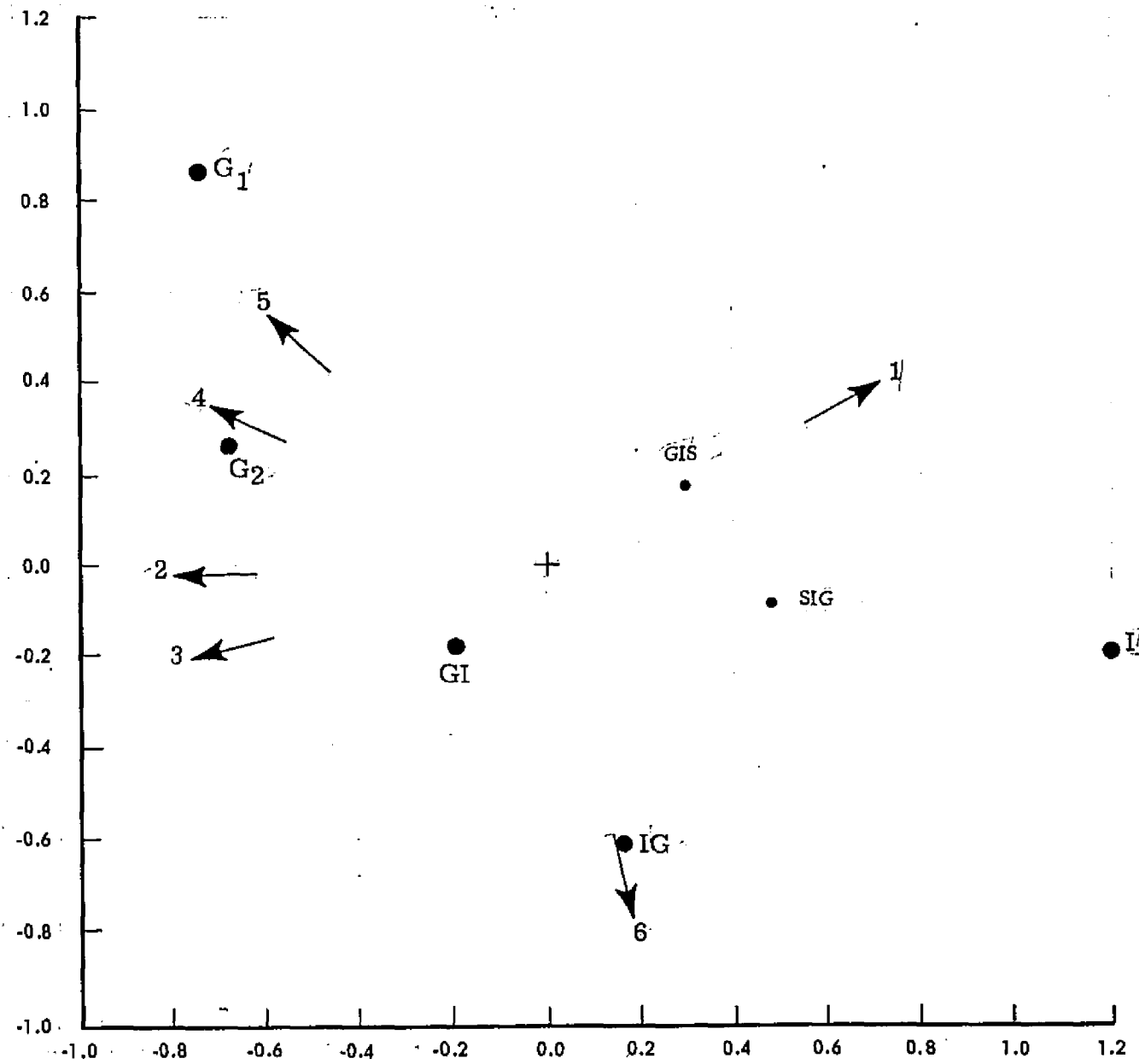
A linear multiple regression procedure was used to map the yield and growth property vectors onto the M-D-SCAL fund solution space of Figure 2. Carroll and Chang's PROFIT (Property Fitting) computer program [7] was used for these analyses. For each property vector, best-fit direction cosines were generated along with a correlation coefficient (see Table A8).

Figure 9 shows the derived property vectors drawn in the cluster centroid plot from the original M-D-SCAL fund space. For the yield property, a best fit vector (number 1) was found to form a positive angle of 29 degrees with the horizontal. Fund point projections on this vector showed excellent correlation (0.930) with the funds' measured yields.

The four property vectors representing growth realizations for the funds over differing periods subsequent to March 31, 1971 (numbers 2, 3, 4 and 5) occur generally together in Figure 9. At least two of these vectors (4 and 5) approach a normal alignment to the income vector as would be expected if growth, in addition to income, was indeed a basic factor accounting for portfolio composition differences.

Figure 9

MUTUAL FUND YIELD AND GROWTH PROPERTY VECTORS
FITTED INTO THE M-D-SCAL SOLUTION SPACE
(Funds Represented By Their Cluster Centroids)



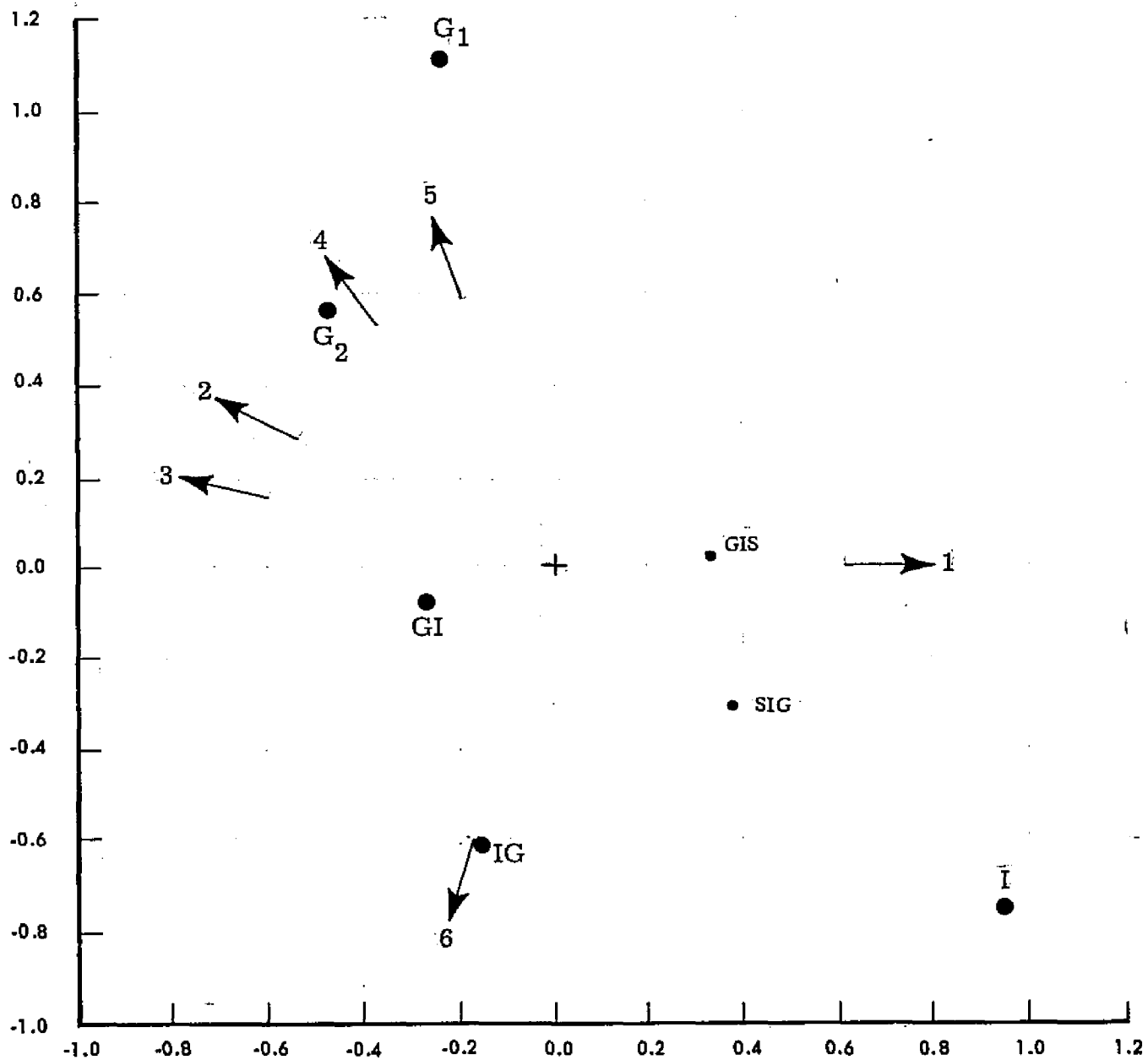
Everything considered, the array of yield and growth property vectors suggest that a more meaningful solution can be achieved if we rotate the reference coordinates orthogonally and anchor the horizontal axes in the direction of the yield property vector. This clockwise rotation of the scaling solution space through an approximate 29 degrees is reflected in the revised cluster centroid plot of Figure 10. In doing this rotation, the earlier seeming discrepancies in cluster positions are resolved. Now, cluster IG has a higher Growth rating than cluster I; also, the relationships of clusters GIS and SIG to cluster GI make more sense than before.

Although the high correlation coefficient between actual fund yields and the fund projections on the horizontal axis in the rotated solution space represents a strong confirmation of an Income label for the abscissa in Figure 10, a label of Growth for the vertical axis cannot be assigned with the same degree of assurance. Correlation coefficients for the four ex post growth property vectors 2, 3, 4 and 5 are .542, .115, .240 and .624, respectively, and thus are unconvincing as to the existence of a growth effect in the mutual fund space.

Undoubtedly, however, the reason for these vexing results is that the chosen property vectors--ex post growth realizations--are inadequate surrogates for ex ante growth expectations. A different route must be taken in trying to measure growth expectations. One possible approach is the hypothesis that growth expectations at any point in time

Figure 10

M-D-SCAL SOLUTION SPACE
ROTATED 29 DEGREES CLOCKWISE
(Compare to Figure 9)



are formulated on the basis of, and can be measured by, historic growth performance.

To test this hypothesis, an additional growth property vector was established corresponding to the historic growth achieved by each fund over the thirty-seven months from December 1, 1968, to March 31, 1971. These growth values again were taken from the Lipper service and also are included in Table A7. This additional property vector was mapped onto the M-D-SCAL fund space using the PROFIT program (see vector 6 in Figures 9 and 10). The historic growth property vector fell in a direction almost exactly opposite of the concluded Growth direction in Figure 10 with a correlation coefficient of .765. This implied the seemingly perverse results that G_1 and G_2 type funds are the least growth oriented!

The failure of these property vector tests to convincingly establish the Growth dimension should not be regarded as a failure in technique. It is just another manifestation of the oft-reported behavior of growth-oriented funds at certain times, under certain business cycle and market conditions to perform worse than mutual funds generally. At the beginning of the last quarter of 1968 Standard & Poor's Composite Stock Price Index was at 102.9. Nearly three years later, on March 31, 1971, the index was nearly identical at 103.3. In between the market dropped to 71.2 and recovered slowly. During this period growth-oriented funds, especially the so-called "go-go" funds (generally

corresponding to the G_1 group, who speculated in high-volatility stock issues) were affected most severely. This led to a widespread retreat by fund managers from the philosophy of striving for extreme growth. Income-oriented funds came through this period in noticeably better shape.

At the close of the four quarters following March 31, 1971, the Standard & Poor's Composite Stock Index stood at 99.7, 98.3, 102.1 and 107.2, respectively. The relative stability in the stock market over this period with an upturn at the end appears to reconcile quite well with the order of arrangement of the four ex post growth property vectors in Figure 10. It appears likely that only in times of general market upswing will ex ante growth expectations show good correlation with achieved ex post growth. Stated in a different way, as many investigators have reported [8, 9, 10, 11, 12], there is little demonstrated evidence that mutual funds on average can outperform or behave differently than the general market or randomly selected unmanaged portfolios. Highly growth-oriented funds have tended to exaggerate market swings, both upward and down.

To summarize the property vector fittings, a mutual fund's current yield can be considered a reliable measure of the fund's emphasis on the Income objective. However, no measure of realized growth can be considered a reliable surrogate for growth expectations at any given point in time. Therefore, it is concluded that the most meaningful

manner of adjusting the fund solution space is to rotate the original orthogonal reference axes to make the Income dimension coincident with the current yield property vector. In doing this, the configuration of mutual fund points, clustered according to their investment objective categories, falls into neat agreement with the labeled coordinate axes.

HIGHER LEVEL FACTORS

At this point persuasive evidence has been marshalled showing that similarity scaling in two dimensions gives interpretable results with the two factors representing income and expected growth investment objectives. Now an attempt must be made to answer whether or not the addition of a third dimension would improve the solution. Can the third factor be interpreted?

In accordance with some of the popular systems used for classifying mutual fund investment objectives (see page 51), a "stability" interpretation of the third dimension immediately comes to mind. To test this conjecture, two groups of three mutual funds each, corresponding to the SIG and GIS investment objective categories, purposely were included in the subject sample to see if any uniqueness in their scaling spatial locations could be explained by a third dimension. If the projections of these funds, especially those belonging to the SIG group, score unusually high on the third factor, support would be lent to a "stability" meaning for this factor.

Column 7 in Table A5 gives the coordinates of the fund points on the third factor in the three-space solution. The range of scores on this factor across the 46 funds of all investment objective types is -1.441 to 0.829 with an average of -.004. Funds 44, 45, and 46 are the SIG funds with coordinate values on the third factor of 0.088, 0.331 and 0.145, respectively, giving an average 0.188. Funds 41, 42, and 43 constitute the GIS group with an average of 0.024 derived from respective values of 0.244, -0.160, and -0.012.

Obviously, there appears to be no special tendency for the so-called stability-related mutual funds to be influenced one way or the other by the third dimension. This evidence does not support a Stability label for factor three.

If the third factor did correspond to a stability dimension, we would expect scale values on this factor (Z) to show a good correlation with the scores on the Income dimension (X). Conversely, if dimension three corresponded to the reciprocal of stability, risk, we would expect the scale values to show a special correlation with those on the Growth dimension (Y). The matrix of correlation coefficients between the three dimensions in Table A5 was computed. Between dimensions X and Z the correlation coefficient was 0.2687; between dimensions Y and Z it was 0.3683. From this neither a stability nor a risk interpretation of the unknown third factor was supported.

Lack of a decernable risk factor in the scaling solution, as discussed in the Introduction, would not disagree with observations by Clarkson and others about how investors really behave. Or, if a volatility-related factor actually exists, it may not be detectable in a study of portfolio holdings at the level of industry groupings. Rather, differences as to perceived risk may only be significant at an individual security and company level.

A label for the third dimension in the scaling solution space has remained elusive in the tests conducted. In line with the stated intent of this study, continued effort to identify the source of variation associated with the third and higher factors is not justified if the higher dimensionality solutions provide only slightly better descriptions and classifications of the original data.

Because it accounts for but a relatively small improvement in configuration stress, the addition of a third factor has little affect on the informational content of the solution pattern. This is shown by the fact that the first two dimensions of the three-space solution are very similar to the corresponding axes of the two dimensional solution space. A correlation coefficient of 0.983 was found in comparing the scale values on the X dimensions in the two- and three-space solutions in Table A5 and a correlation coefficient of 0.882 was found in comparing their respective Y dimensions.

Also, to test the information contribution of the third factor, a HICLUST analysis was performed on the M-D-SCAL three dimensional solution as it was for the two dimensional solution. Figure A6 (b) gives the dendritic HICLUST output for the three-space configuration. The hierarchy of natural cluster groupings is very much the same as the two-space results shown in Figure A6(a) and plotted in Figures 6 and 7. Maximum likelihood probabilities for the natural cluster labels were computed for the three-space display as follows:

Table 6

DEGREE OF CORRESPONDENCE BETWEEN
NATURAL AND PREDICTED CLUSTERING
FOR THREE DIMENSIONAL M-D-SCAL SOLUTION

Natural Cluster	Maximum Likelihood Investment Objectives	Maximum Likelihood Probability
<u>Primary Hierarchical Level[See Figure A6 (b), row 2]</u>		
Cluster 1	G (inc. GIS)	0.675
Cluster 2	I (inc. SIG)	1.000
Overall		0.825
<u>Secondary Hierarchical Level[See Figure A6 (b), row 4]</u>		
Cluster 1	G (G ₁)	0.556
Cluster 2	GI (inc. G ₂ & GIS)	0.667
Cluster 3	IG (inc. SIG)	0.474
Cluster 4	I	1.000
Overall		0.609

Comparing these results with Table 5 shows that the three-space solution is no better and perhaps slightly worse than the two-space solution, at least in the sense of the natural clusters representing prior-defined categories of mutual fund investment objectives.

Synopsis

This chapter has demonstrated the ability of multidimensional scaling to meet the stated research goals of the study; namely, to find an interpretable factor structure of portfolio evaluation and, simultaneously, to scale existing portfolios with respect to the factors uncovered. In accomplishing this, a new level of information becomes available to prospective mutual fund investors. Despite the categorical nature of conventional fund classification schemes, mutual funds simply do not fall into completely homogeneous groups. Rather, they cover a continuous graduated range of emphasis on investment objectives such as income and growth. Remarkably, the nonmetric multidimensional scaling techniques are able to reconstruct this continuum of investment objectives in a spatial display of mutual funds using only the information contained in a table of the funds' portfolio holdings. No external information is necessary as to past performance or predicted future performance of the funds or their component assets.

Everything considered, for the descriptive and classificatory purposes of this study, a two dimensional solution space is sufficient. In spite of a better fit to the original data, going from a two-space to a

three-space display barely alters the informational content of the solution and tends to diminish its comprehensibility. For these reasons, the two-space configuration is presented as the primary result of this study.

The research techniques presented in this study cannot properly be regarded as useful tools for predicting the growth performance of individual mutual funds. Evidence shows convincingly that only under certain market conditions do growth expectations correlate well with subsequent growth realizations. It remains for the investor, himself, to predict future economic and business conditions and to decide for his circumstances a favored trade-off between Growth versus Income emphasis. However, once a business cycle scenario is selected, the scaling display offers to show which current portfolios best meet the investor's chosen combination of investment objectives. Any Growth/Income trade-off can be portrayed as a family of convex iso-preference, or indifference curves in the two dimensional solution space. Therefore, the real usefulness of the scaling methods presented in this study is to show at any given point in time, for any favored combination of investment objectives, which mutual fund(s) is optimal.

The foregoing suggests an operational use for the scaling techniques. That is, the techniques would be implemented as an ongoing monitor to detect changes in investment objectives for individual mutual funds and, indeed, for the securities market as whole. The next chapter investigates the sensitivity of the experimental results to changes over time.

REFERENCES

Chapter 3

1. J. Kruskal and F. Carmone, "How to Use M-D-SCAL, A Program to Do Multidimensional Scaling and Multidimensional Unfolding (Version 5M of MDSCAL all in Fortran IV)," mimeographed report (Bell Telephone Laboratories, Murray Hill, N.J., October, 1969).
2. P. Green and V. Rao, Applied Multidimensional Scaling: A Comparison of Approaches and Algorithms (New York: Holt, Rinehart and Winston, Inc., 1972), pp. 188-192.
3. Forrest W. Young, "TORSCA, an IBM Program for Nonmetric Multidimensional Scaling," Journal of Marketing Research, Vol. V (August, 1968), pp. 319-321.
4. Forrest W. Young, "Bibliography of Multidimensional Scaling Programs Developed at the University of Michigan, Bell Telephone Laboratories and the University of North Carolina Including Condensed Descriptions and Abstracts of Unpublished Programs," mimeographed paper prepared for the Bell-Penn Workshop on Multidimensional Scaling (Philadelphia: University of Pennsylvania, June 7-10, 1972).
5. Stephen C. Johnson, "Hierarchical Clustering Schemes," Psychometrika, Vol. XXXII (September, 1967), pp. 241-254.
6. Arthur Lipper Corporation, "Mutual Fund Performance Analysis," a weekly bulletin (New York, March 31, 1971).
7. J. J. Chang and J. D. Carroll, "How to Use PROFIT, a Computer Program for Multidimensional Analysis of Preference Data," mimeographed report (Bell Telephone Laboratories, Murray Hill, N.J., 1969).
8. Irwin Friend, et al, "A Study of Mutual Funds," prepared for the SEC Report of the House Committee on Interstate Commerce & Finance, 87th Congress, 2nd Session (U. S. Government Printing Office, 1962).
9. I. Friend, M. Blume and J. Crockett, "Mutual Funds and Other Institutional Investors: A New Perspective," A Twentieth Century Fund Study, 1970.
10. Michael D. Jensen, "The Performance of Mutual Funds in the Period 1945-1964," Journal of Finance, Vol. XXIII (May, 1968), pp. 389-416.

Chapter 4

SCENARIO DEPENDENCE OF INVESTMENT PREFERENCES

The experiment of the last chapter revealed a factor structure of portfolio composition for a single point in time. A natural line of inquiry from both financial and statistical standpoints concerns the responsiveness of scaling results to portfolio changes over time. Hopefully, the experimental methods would be sufficiently sensitive to measure shifts in individual mutual fund investment objectives as reflected in adjustments to their portfolio mixes. It is reasonable to believe that as time progresses, a fund point will not remain fixed but will migrate in the Growth/Income investment objective space as fund managers respond to unfolding economic conditions.

This chapter reruns the similarity multidimensional scaling experiments with new data representing mutual fund holdings over a sequence of short-term time intervals. Last chapter's analysis of portfolio holdings for March 31, 1971, has been repeated for the remaining three end-of-quarters of 1971, i.e., June 30, September 30 and December 31. Computational considerations constrained the analysis to four points in time. The choice of particular dates--the concentration on short-interval portfolio changes--was dictated by data availability. However, data availability aside, focusing on short-term effects was purposeful. Any real usefulness of the experimental methods

from this study would rest in their ability to function in an on-going quick-response manner--to analyze information as frequently as it becomes available so as to give early indication of changes in fund investment strategies and shifts in the overall market outlook.

Data Preparation and Similarity Scaling

With the original sample of 46 mutual fund subjects held intact, the data preparation steps of chapter 2 and the similarity scaling experiment of chapter 3 were repeated for the last three quarters of 1971. For each end-of-quarter reporting date, mutual fund portfolio holdings were obtained from Arthur Lipper's INSIGHT publication and were aggregated into the 97 standard industry and asset groups. Altogether, the portfolio assets data for all funds over the four quarters represented approximately 18,000 bits of data.

The three new 97 x 46 stimuli-by-subjects preference matrices were transformed into three 46 x 46 subjects-by-subjects holdings overlap similarity matrices. Each new proximity matrix was subjected, in turn, to an M-D-SCAL analysis over a range of dimensionalities. Stress results for each of the new three quarterly solutions in one to five dimensions were found to be nearly identical to the corresponding stress for the First Quarter solution in Table 3. Tables A8, A9, and A10 give the resulting Euclidean space configurations for the two and three dimensional solutions. Figures 11, 12 and 13 show the key two-space displays (reflected around the X axis to have increasing Growth in the upward

Figure 11

M-D-SCAL MUTUAL FUND SOLUTION SPACE--JUNE 30,
WITH ENCLOSED INVESTMENT OBJECTIVE GROUPS

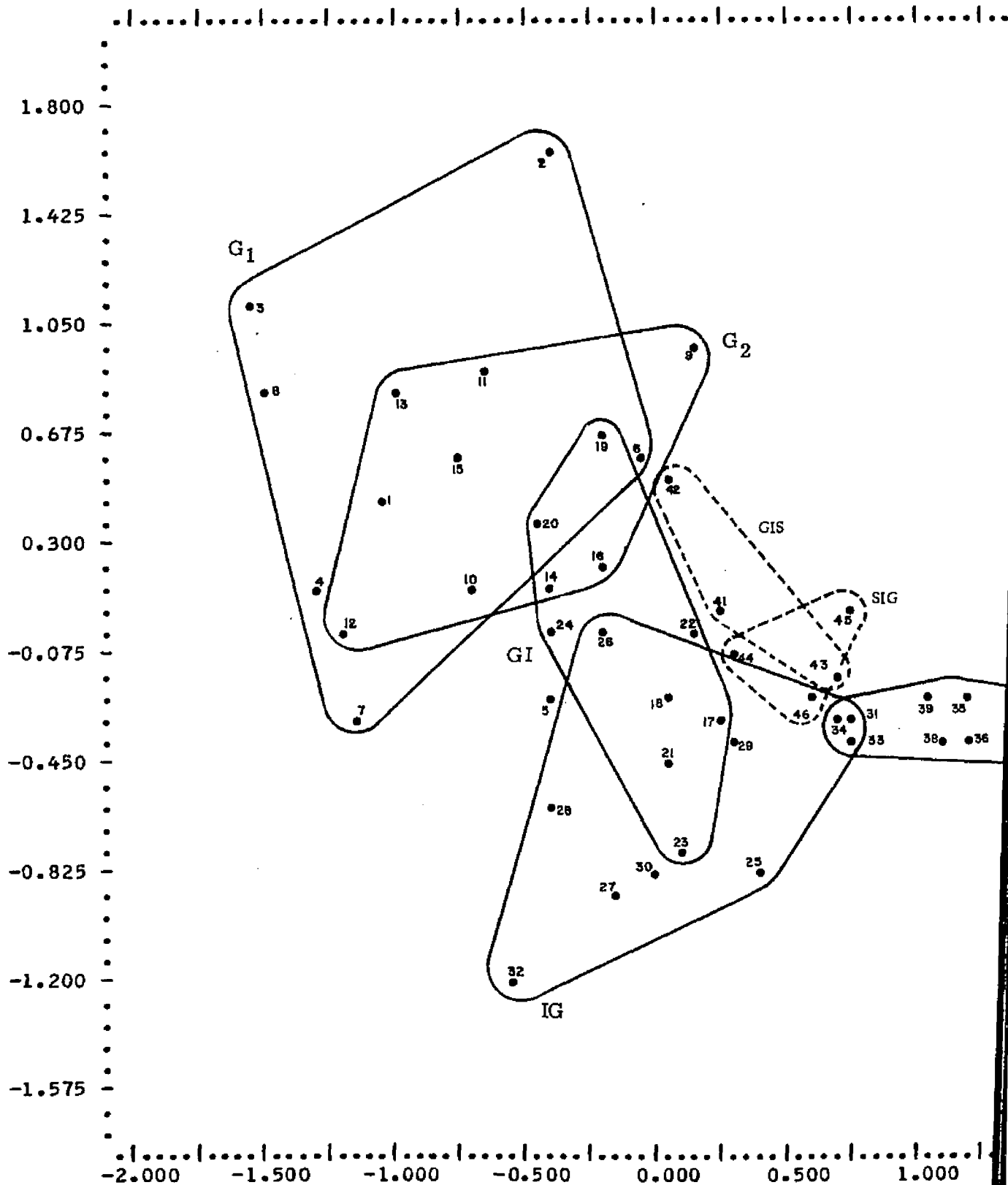


Figure 11

D-SCAL MUTUAL FUND SOLUTION SPACE--JUNE 30, 1971
WITH ENCLOSED INVESTMENT OBJECTIVE GROUPS

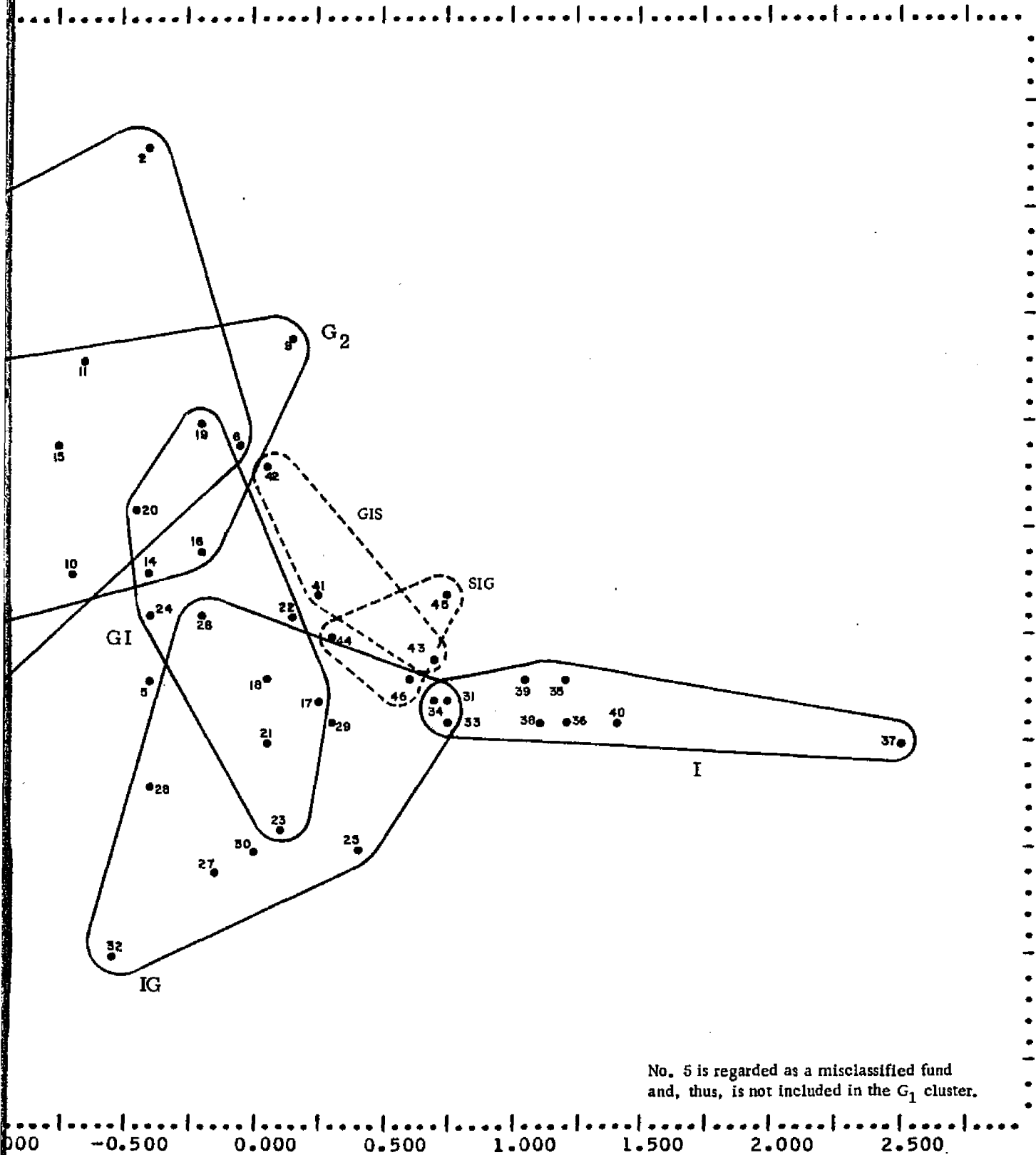


Figure 12

M-D-SCAL MUTUAL FUND SOLUTION SPACE--SPETEMBER
WITH ENCLOSED INVESTMENT OBJECTIVE GROUPS

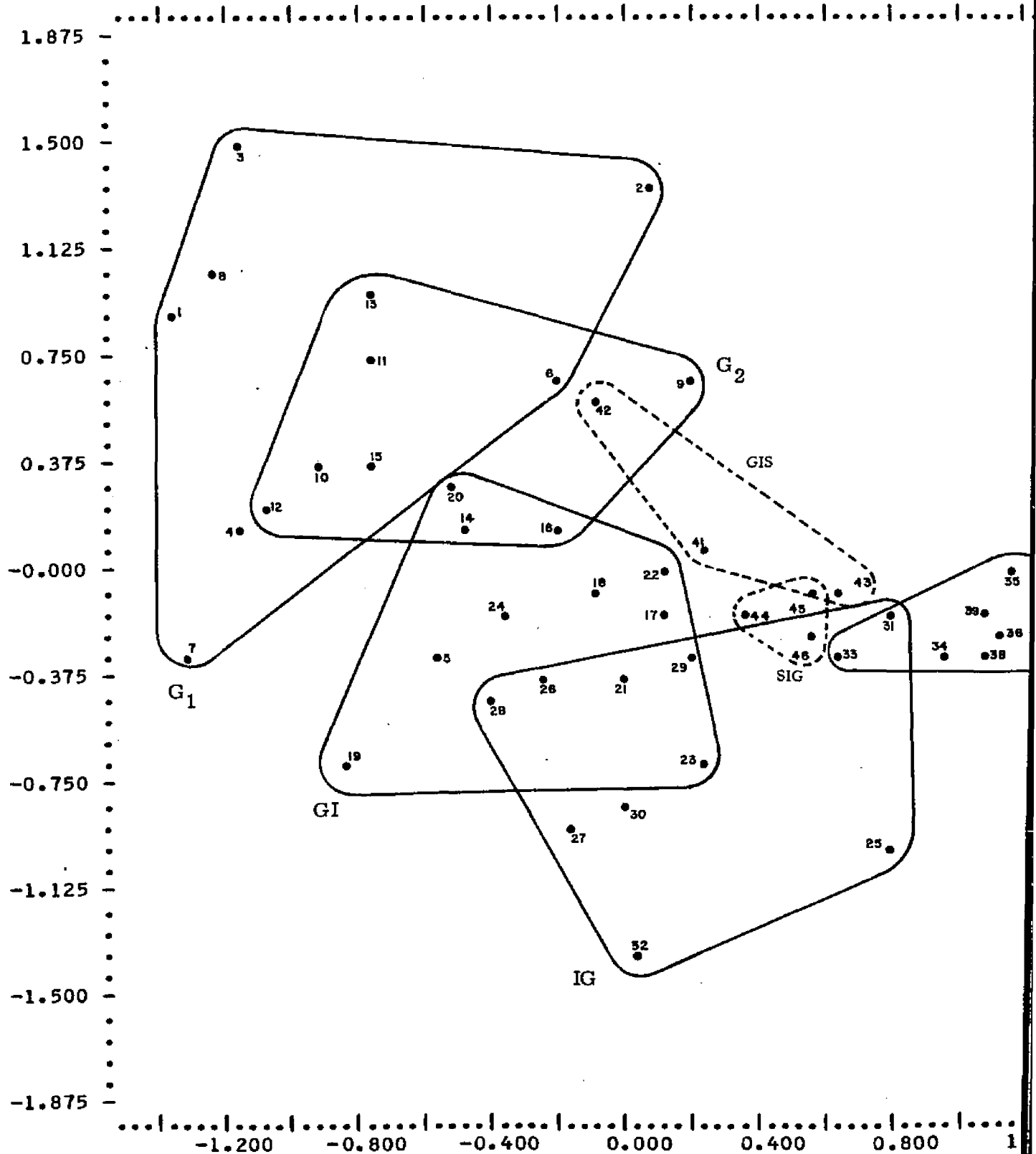
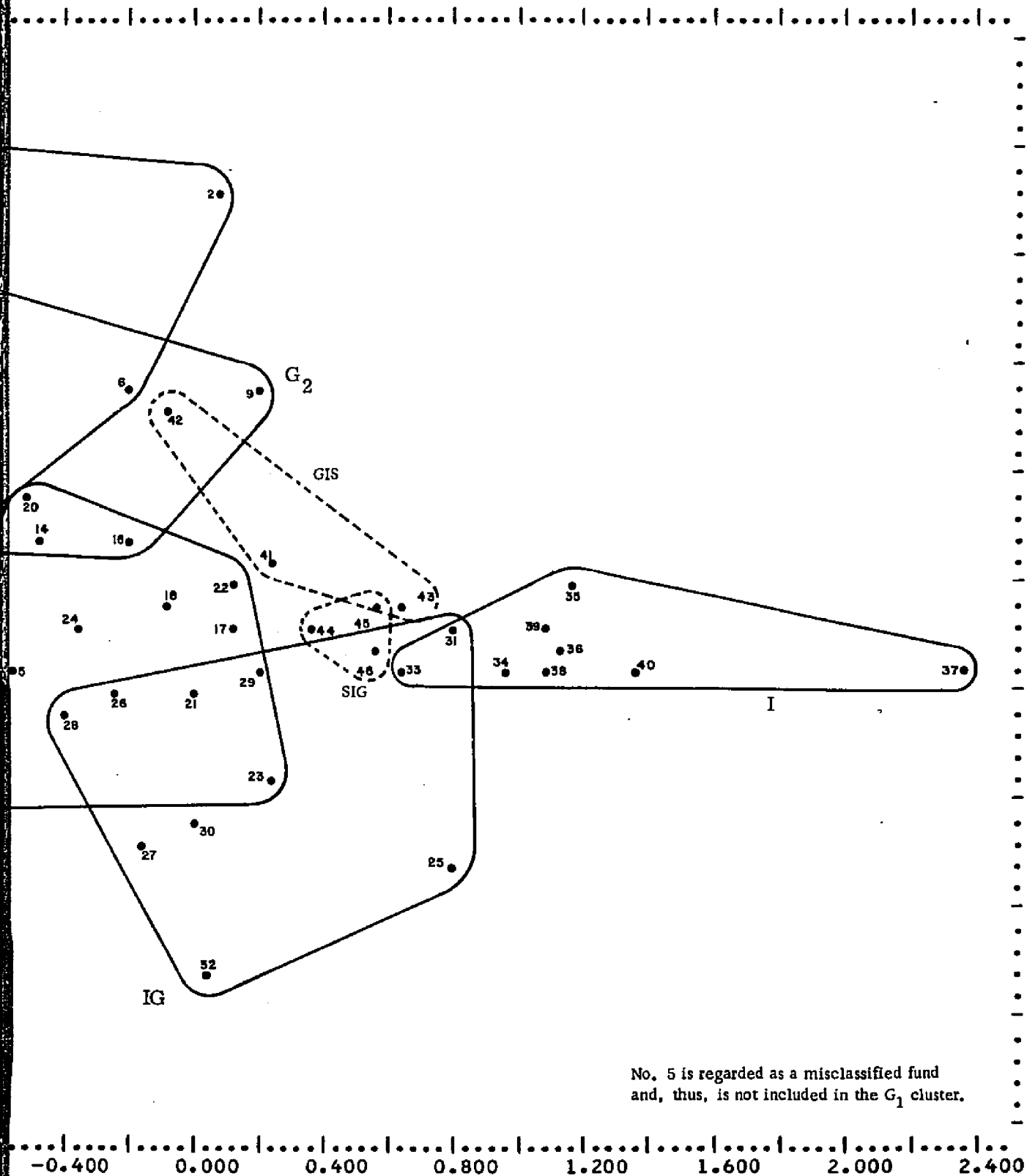


Figure 12

ALL MUTUAL FUND SOLUTION SPACE--SEPTEMBER 30, 1971
WITH ENCLOSED INVESTMENT OBJECTIVE GROUPS



No. 5 is regarded as a misclassified fund
and, thus, is not included in the G₁ cluster.

Figure 13

M-D-SCAL MUTUAL FUND SOLUTION SPACE--DECEMBER
WITH ENCLOSED INVESTMENT OBJECTIVE GROUPS

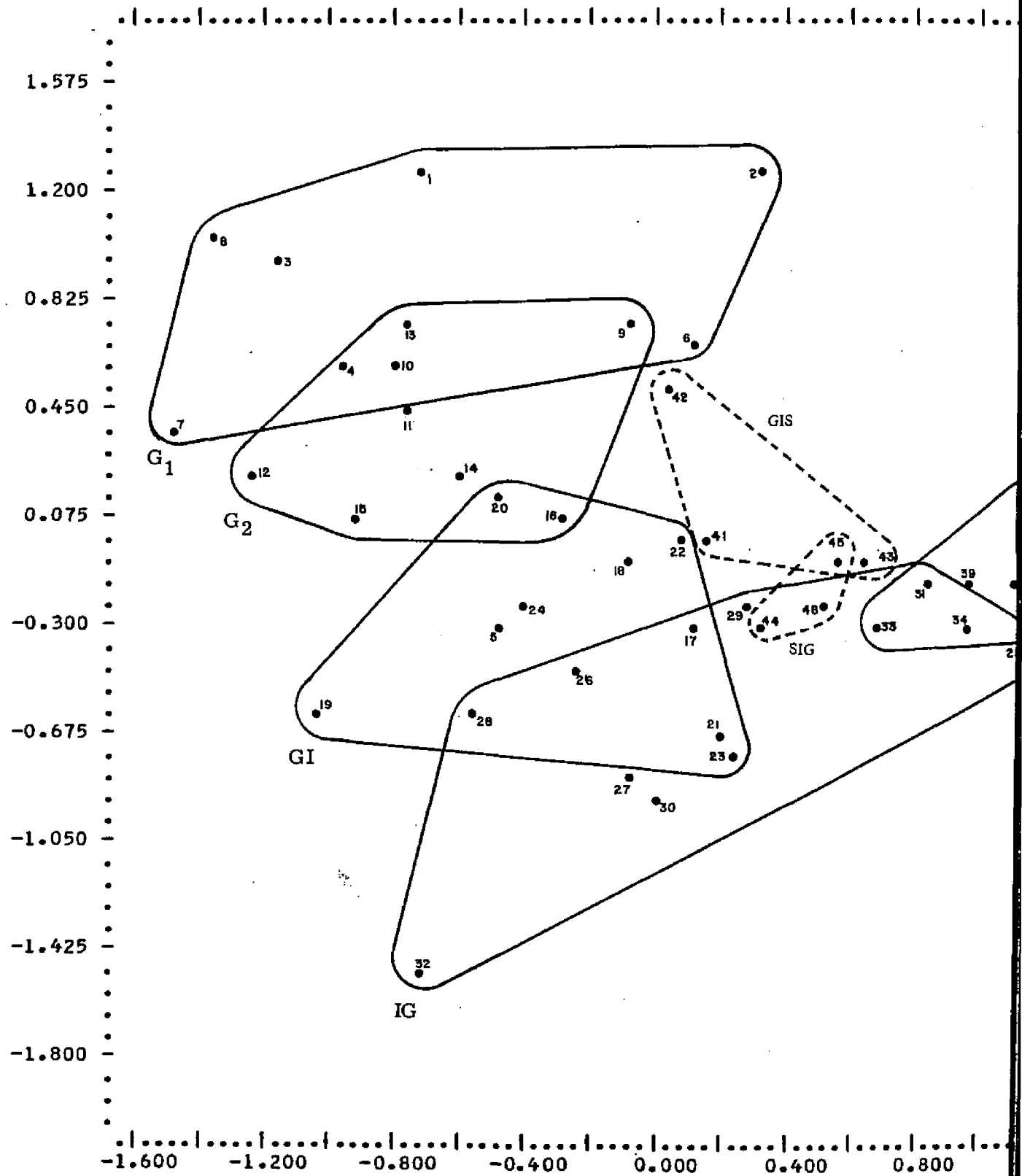
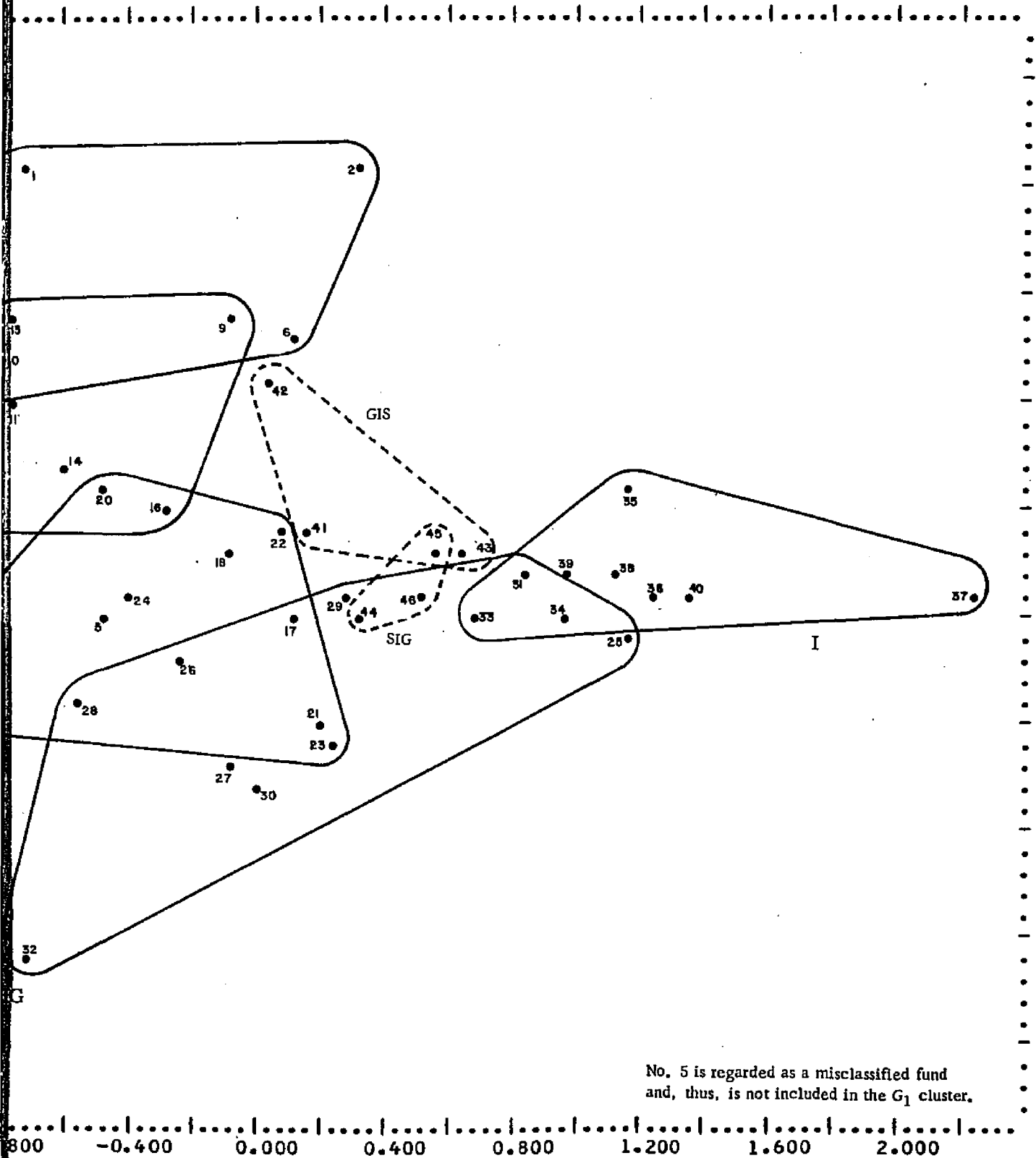


Figure 13

CAL MUTUAL FUND SOLUTION SPACE--DECEMBER 31, 1971
WITH ENCLOSED INVESTMENT OBJECTIVE GROUPS



direction). Within each have been drawn a priori cluster contours to define convex sets of funds of the same investment objective class. The set of four solution displays, Figures 5, 11, 12 and 13, can be regarded as a series of snapshots showing the mutual fund evaluative space as it changed through 1971.

Visual comparison of the separate end-of-quarter spaces, including that for the First Quarter in Figure 5 shows a substantial degree of constancy in the configurations. Clusters retain basically the same overlapping relationships to each other. At the level of individual elements, however, some mutual fund points show sizable movements from quarter to quarter. For example, see the change in positions of funds 7 and 9 from the First Quarter to the Second. The portion of change in fund 7's position in the direction of less Income can be traced to the fund's reduction of long-term debt assets over this period (16.9% of total assets to 2.3%).

CONFIGURATION MATCHING

Instead of discussing the time-sensitivity of scaling results by reporting the movement of individual fund points, it would be useful to be able to summarize in some single measure the relationship between entire configurations of points representing different instants in time. Several methods are available for measuring the degree of correspondence or congruence of different sets of points in some space; namely -

1) Product moment correlation--

A configuration of points can be translated into a vector of ratio-scaled data values representing all interpoint distances within the space. Two different configurations can be compared by computing the correlation coefficient between their respective vectors of interpoint distances. Comparisons between many configurations would result in a matrix of pairwise correlation coefficients. Such correlation coefficients are invariant under similarity transformations of the sets of points.

2) Cannonical correlation--

Coordinates of one scaling solution can serve as the criterion set in a conventional cannonical correlation analysis and coordinates of a second scaling solution can serve as the predictor set. The results of this analysis would be cannonical weights and a cannonical correlation coefficient. This coefficient would measure the correspondence between the configuration allowing for translation, reflection, rotation and differential stretching or compression of the axes of either space.

The correlation indices associated with both of the techniques above are useful for summarizing how well configurations agree with each other on a pairwise basis. However, it would be better if a method could be devised to allow the simultaneous comparison of many solution spaces. And unlike correlation coefficient, the measure of comparison should have meaning in terms of the basic factors which describe the spaces.

A method of comparing configurations which meets these requirements is available in Carroll and Chang's INDSCAL (Individual Differences Scaling) computer program. The configuration matching option of this program is capable of simultaneously measuring the congruence of many separate sets of points against some fixed target

configuration. This target or reference configuration is input to the program as coordinates of its member points. Each trial configuration to be matched against the target is reduced to a matrix of its respective interpoint distances and, as such, also is entered. The output from the INDSCAL program are two separate spatial displays spanned by the same set of dimensions. They are on one hand, the original target configuration and, on the other, a second-order configuration of pseudo-subject points, each point standing for one entire trial configuration.

The coordinates for any one of these pseudo-subject points represent the respective trial configuration's idiosyncratic weightings or saliences on the dimensions of the target space. Thus, the trial configuration as represented by a particular pseudo-subject point can be best reconstructed from the target configuration by differentially stretching or shrinking the target configuration's dimensions in proportion to the square roots of the respective pseudo-subject point's dimension weights. The INDSCAL program produces these dimension weights and, in addition, generates a product-moment correlation coefficient between the trial configuration and the transformed target space.¹

1. When the dimensions of the target space are normalized (so that the sum of square coordinates on each dimension equals one), the squared weight indicates the proportion of variance in the trial configuration which can be accounted for by that dimension. The square of a weight underestimates the proportion of variance accounted for if, as in the present study, the dimensions of the reference space are correlated (0.451). See Wish and Carroll 1 .

An INDSCAL configuration matching analysis was performed on our set of four quarterly M-D-SCAL solution spaces in two dimensions. The First Quarter solution space was taken as the fixed reference configuration. Interpoint distances within each of the four quarterly solution spaces were computed (including that for the First Quarter as a check) and were treated as trial configurations. Table 7 below gives the INDSCAL results, the dimension weights for each trial configuration as well as its goodness of fit correlation to the respectively transformed target space.

Table 7

<u>Pseudo-subject</u>	<u>Dimension Weight</u>		<u>Goodness of Fit</u>
<u>End-of-Quarter</u>	<u>Income</u>	<u>Growth</u>	<u>Correlation</u>
1 (March 31, 1971)	0.826	0.431	1.000
2 (June 30, 1971)	0.776	0.316	0.890
3 (Sept. 30, 1971)	0.777	0.364	0.917
4 (Dec. 31, 1971)	0.764	0.335	0.889

Scenario Dependence

The ability of INDSCAL to characterize entire solution spaces by a set of coordinates or weights on the basic factors of investment evaluation is one of the most promising results of the present study. The idea of investment objective weights for the overall market is a meaningful notion. At any point in time the market for securities can be said to take on an average or consensual direction with respect to the basic factors of investment preference. That is to say the market

at any point in time is so much (X) Income oriented and so much (Y) Growth oriented. The INDSCAL analysis can assign interval scaled numbers to these X and Y investment objective weights, at least in a sense relative to some reference market condition. What is suggested is that INDSCAL dimension weights may serve as a new form of barometer for overall market attitudes in terms of current Growth and Income emphasis. Present market indicators such as the Dow Jones or Standard & Poor's composite stock price indices are neutral as to market expectations. To the extent that the investment "psychology" of a large sample of mutual funds mirrors the market attitude generally, a barometer based upon such configuration matching dimension weights could be useful in research relating economic conditions (scenarios) to market performance. Thus, the dimension weights may have some predictive value, at least for forecasting the direction of the market as a whole if not for individual funds.

With the small sample of market changes in the present study (three intervals), the conjecture of predictive value of INDSCAL dimension weights could not be tested rigorously. However, the small differences between the dimension weights of each end-of-quarter do not conflict with the fact that 1971 was a period of relatively stable end-of-quarter stock prices.

The small differences between dimension weights in Table 7 are to be expected. From quarter to quarter the portfolio of a typical

fund remains largely intact. The turnover rate for the common stock portion of investment company portfolios was found to average about 14 percent per quarter [2].

Obviously, the only manifestation of shifts in investment objectives are the changes in portfolio composition. This suggests that changes in dimension weights for individual funds as well as for the market as a whole can be magnified for closer examination by an experimental design which works only with quarterly changes in portfolio holdings, not with the absolute holdings themselves. The use of portfolio change data in this manner represents a possible improvement in experimental design and remains an open avenue for future investigation.

CANNONICAL DECOMPOSITION

Before leaving the topic of scenario dependence of investment preferences, an alternative means of using the INDSCAL computer program for this type of analysis should be mentioned. The configuration matching option of INDSCAL discussed above, operates on a three-way scenario-by-subject-by-subject matrix of interpoint distances. These interpoint distances derive from the scaling solution configurations of separate M-D-SCAL analyses of subject-by-subject holdings overlap similarities matrices representing different points in time. The holdings overlap matrices, in turn, are computed from separate subject-by-stimulus matrices of investment group preferences. The canonical decomposition (CANDECOMP) option of INDSCAL sidesteps

all these intermediate similarities scaling steps and operates directly on one scenario-by-subject-by-stimulus preferences matrix. For any prespecified dimensionality the program simultaneously produces the following three "joint" spatial solutions, each spanned by the same system of dimensions:

- 1) A time-and-mutual fund-averaged stimuli space of industry and asset group points
- 2) A time-averaged subject space of mutual fund points
- 3) A pseudo-subject space of time (scenario) points

The INDSCAL program treats data meterically and offers the valuable feature of giving a solution space of unique orientation.¹

Considerable effort was expended investigating the applicability of CANDECOMP to the present study. The three-way matrix of holdings preferences were analyzed in two dimensions. The mutual fund subject configuration in the resulting joint space display showed some semblance to the M-D-SCAL solution spaces. However, the industry group stimuli space and the scenario pseudo-subject space were not visually interpretable. For reasons discussed in the next chapter under the heading, "Joint Space Models--Preference Scaling," exploration of the CANDECOMP

1. See Green and Rao for a discussion of canonical decomposition of scenario dependent preferences [3]. They advise that the M-D-SCAL/INDSCAL configuration matching route is possibly a superior approach than the direct CANDECOMP analysis.

approach was terminated. Nonetheless, the use of INDSCAL in this manner perhaps is worthy of further study.¹

1. Another research tool essentially identical to INDSCAL is Richard Harshman's program PARAFAC 1 [4]. PARAFAC 2 is a different model with certain features that perhaps make it a more suitable tool. See the discussion of system variation in Harshman [5].

REFERENCES

Chapter 4

1. M. Wish and J. D. Carroll, "Applications of 'INDSCAL' to Studies of Human Perception and Judgment," mimeographed paper, (Bell Telephone Laboratories, Murray Hill, N.J., 1971), to be published in Handbook of Perception, ed. M. P. Friedman, (New York: Academic Press, 1973).
2. United States Securities and Exchange Commission, "Institutional Investor Study Report of the Securities and Exchange Commission," report to the Committee on Interstate and Foreign Commerce, 92nd Congress, 1st Session (U. S. Government Printing Office, 1971), Vol. VIII, p. 22.
3. P. Green and V. Rao, Applied Multidimensional Scaling: A Comparison of Approaches and Algorithms (New York: Holt, Rinehart and Winston, Inc., 1972), pp. 101-103.
4. Richard A. Harshman, "PARAFAC," Working Papers in Phonetics No. 16 (University of California, Los Angeles, December, 1970).
5. Richard A. Harshman, "PARAFAC 2," Working Papers in Phonetics No. 22 (University of California, Los Angeles, March, 1972), pp. 7-9, 30-44.

Chapter 5

EXPERIMENT DESIGN VARIATIONS

Previous chapters have referred to several design variations that have been investigated in the course of reaching a final experimental approach. To complete the coverage of the avenues explored in the present study, the following topics are discussed in this chapter:

- 1) Alternative Proximity Measures
- 2) Non-Euclidean Solution Spaces
- 3) Preference Scaling--Joint Space Solutions

Each of these issues represents a major area of methodological variation in the art and science of multidimensional scaling.

ALTERNATIVE PROXIMITY MEASURES

Numerous measures have been reported in the literature for computing pairwise proximities between subjects across their stimuli response profiles.¹ New formulas readily can be designed to suit the particular conditions of a given problem. Holdings overlap, the measure of portfolio similarity in the present study was specially formulated to incorporate certain features stemming from our understanding of actual investment selection behavior. Nevertheless, scaling experiments were repeated using alternative proximity formulas and the

1. For reviews of such measures, see Cole [1], Coombs [2], Green and Rao [3], and Sokal and Sneath [4].

resulting stress levels and interpretability of the solutions were compared. Following are three proximity formulas which were chosen as plausible alternatives to holdings overlap:

- 1) Euclidean Distance in Original Industry Group Space

$$D_{ab} = \sqrt{\sum (h_{ai} - h_{bi})^2}$$

- 2) City-Block Distance in Original Industry Group Space

$$D_{ab} = \sum |h_{ai} - h_{bi}|$$

- 3) Ratio-Overlap¹

$$S_{ab} = \frac{\sum \min (h_{ai}, h_{bi})}{\sum \max (h_{ai}, h_{bi})}$$

Here, D_{ab} designates the formula to be a dissimilarity measure between subjects a and b; S_{ab} indicates the formula to be a similarity measure. h_{ai} and h_{bi} are investor a and b's respective percentages of their total portfolios represented by their holdings in industry group i.

Triangular matrices of paired subjects proximities were computed with each of the formulas from the First Quarter h_i data in Table A3. Separate M-D-SCAL analyses of these proximity matrices were performed over a range of dimensionalities. Scaling experiments based upon formulas (2) and (3) above showed configurations of higher stress than the comparable results using holdings overlap (see Table 8).

1. Used by Gregson [5].

Moreover, the solution configurations were inferior distortions of the holdings overlap results from an interpretability standpoint.

The first of the alternative formulas above, Euclidean distance, is undoubtedly the most popular measure for deriving proximities from multivariate data. Generally this formula is chosen when conditions do not favor any one above others. The formula is equivalent to the Pythagorean theorem in multivariate stimuli (h_i) space and is equivalent to Farrar's goodness of fit index discussed on page 29.

The proximity matrix based upon the Euclidean distance formula was subjected to M-D-SCAL analysis giving the comparative stress results shown in column 3 of the following Table:

Table 8

Stress Versus Dimensionality for Scaling
Solutions Based on Alternative Proximity Measures

<u>Dimensionality</u>	<u>Holdings Overlap Similarity</u>	<u>Euclidean Dissimilarities</u>	<u>City-Block Dissimilarities</u>	<u>Ratio-Overlap Similarities</u>
1	0.514	0.214	0.515	0.529
2	0.400	0.171	0.402	0.421
3	0.325	0.122	0.325	0.338
4	0.269	0.089	0.271	0.296

Dramatic stress improvement over the standard holdings overlap results is seen for the Euclidean distance case. By itself, this evidence forcefully supports the use of the Euclidean formula for computing mutual fund portfolio proximities. However, the visual comprehend-

ability of the Euclidean proximity solution spaces are disappointing. Figure 14 shows the Euclidean proximity two-space display. Recognizable clusters simply do not emerge. Instead, the configuration is closely packed with a notable trend only in the horizontal direction.

A perplexing situation now must be resolved. Does the sharply lower stress reading mean that we must accept the Euclidean proximity solution configuration as a better representation of the true spatial configuration despite it being clearly an inferior reproduction of expected mutual fund groupings? Intuition favors the holdings overlap solution space, if for no other reason that the clearly recognizable clusters it shows are so unlikely to occur with an incorrect procedure. To accept this choice requires an explanation and refutation of the lower stress-value Euclidean proximity results.

An explanation of the low stress values obtained with the Euclidean formula can be traced to the nature of the original industry group preference data. An examination of the investment profile values in Table A3 shows that the typical fund has zero holdings in the majority of the 97 industry and asset groups (the average fund held securities in approximately one-third of the groups). The use of these zeros as preference scale values would be equivalent to the assumption that the investor (mutual fund) had an equal low degree of preference for all industry groups in which it had zero holdings. But this is clearly erroneous. As discussed in the Introduction, because of the effect of the investor's

Figure 14

M-D-SCAL MUTUAL FUND SOLUTION SPACE
BASED UPON EUCLIDEAN DISSIMILARITIES

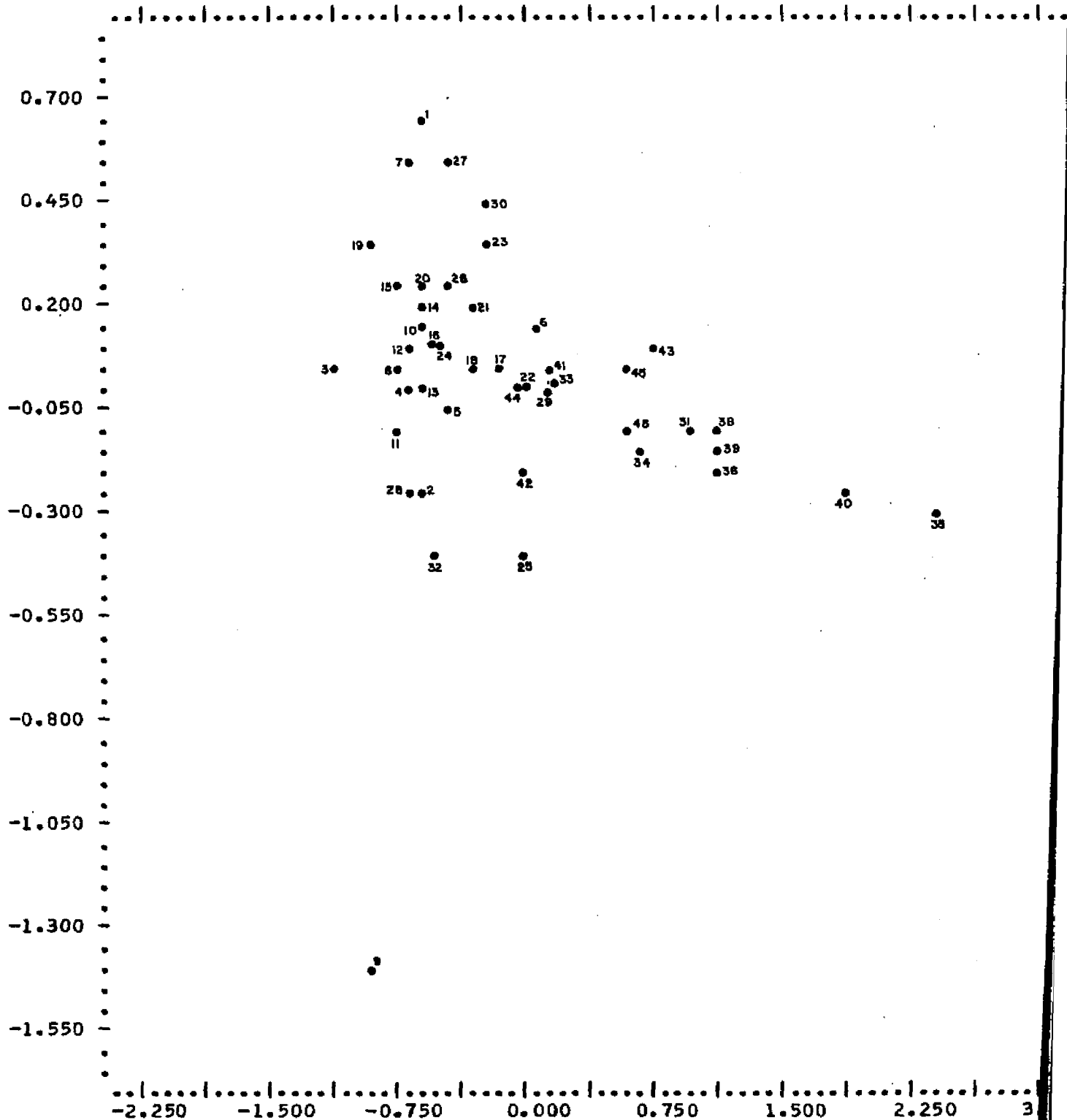
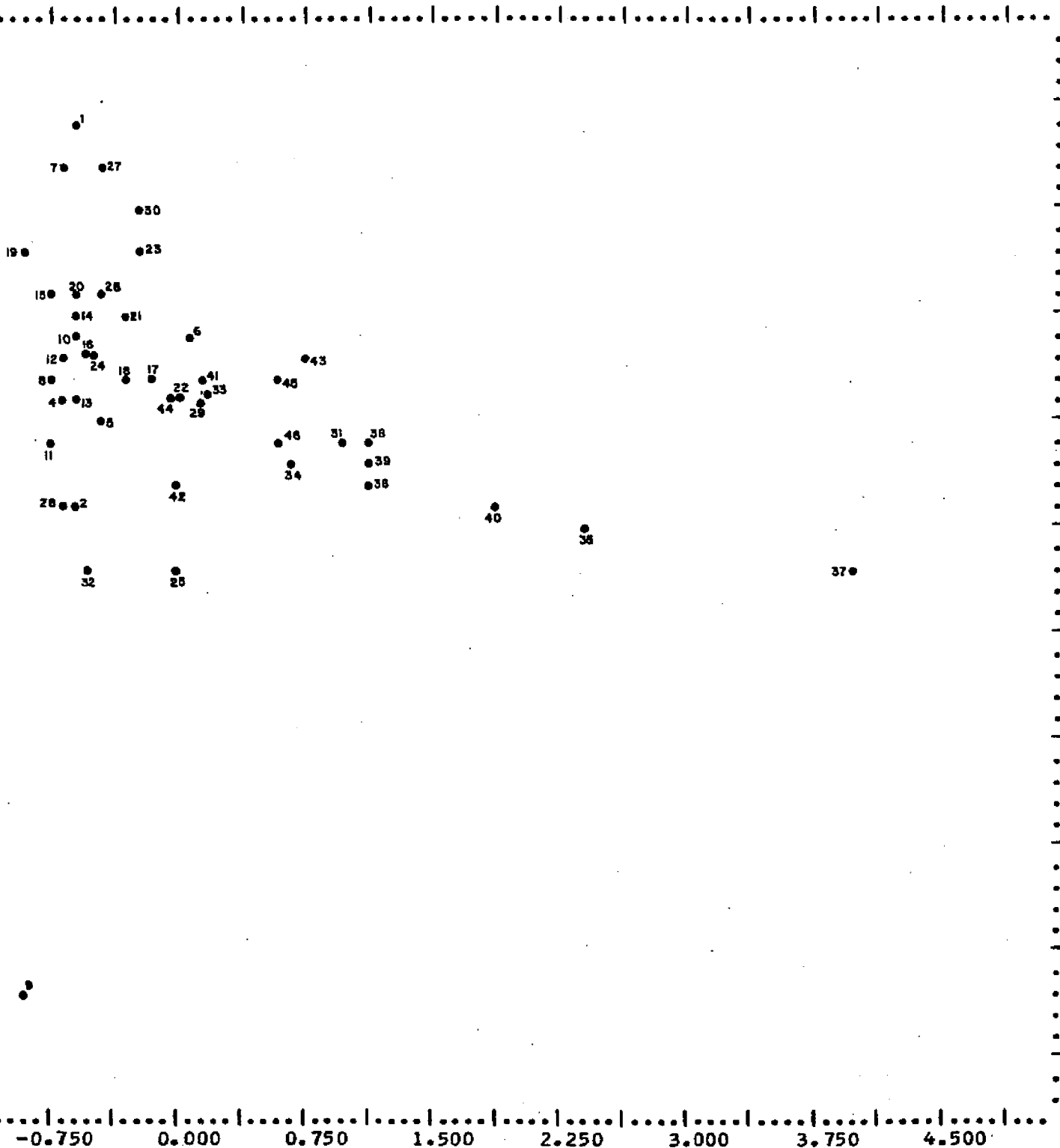


Figure 14

M-D-SCAL MUTUAL FUND SOLUTION SPACE
BASED UPON EUCLIDEAN DISSIMILARITIES



budget constraint and because investors tend to evaluate securities sequentially, an industry group may be unrepresented in an investor's portfolio even if it is a potentially favored group. Zero holdings of industry groups, therefore, are unreliable indicators of the rank order of preferences for the groups.

In this light, pairwise comparisons of mutual funds should be based upon the mutuality of their positive selections of industry groups (non-zero holdings) rather than their similarity in not selecting particular industry groups (zero holdings). In other words, for the portfolio selection problem, the measure of association should exclude similarities in negative attributes and only consider similarities in positive attributes.

These ideas have been expressed in another context. Sokal and Sneath have designed a taxonomy of coefficients of association or resemblance primarily relating to biological and ecological sciences. A fundamental dichotomy in their classification system is whether or not the measure of association takes into account negative attributes. Sokal and Sneath have commented on this point in a biological context as follows:

"Whether negative matches should be incorporated into a coefficient of association may occasion some doubts. It may be argued that basing similarity between two species on the mutual absence of a certain character is improper. The absence of wings, when observed among a group of distantly related organisms (such as a camel, louse and nematode) would surely be an absurd indication of affinity ... it is not pertinent to count 'absence of feathers' when comparing two

bacteria ... Through reducto absurdum we can arrive at a universe of negative character matches purporting to establish the similarity between [any] two entities." [6]

With these arguments in mind, the specification for a measure of portfolio association was set that only non-zero industry group holdings could influence the spatial proximity of points between paired funds in a scaling solution. From this specification, the holdings overlap formula is an immediate consequence; namely -

$$S_{ab} = \sum \min (h_{ai}, h_{bi})$$

With this formula, a contribution is made to similarity, S_{ab} , only if the holdings in investment group i of both funds a and b are non-zero. A zero holding for either fund will result in a zero addition to their computed similarity.

For any scaling subject space which is based upon a similarity measure of proximity such as holdings overlap, the starting condition is a set of points that are maximally separated from each other (zero similarities). Using the overlap formula for any two funds, as industry group holdings are compared, fund points are brought closer together only when both funds have non-zero values in the same group. If one or both of a pair of funds has a zero position in a particular industry group, the effect of the holdings overlap calculation is to leave the respective fund points in unchanged separation.

Dissimilarity measures of proximity such as the Euclidean distance formula work just the opposite. The starting condition is a set of fund points all spatially coincidental (zero dissimilarities). As industry groups are compared, if two funds simultaneously have a zero holding, the effect will be to leave the fund points together.

In short, spatial nearness in an overlap similarities solution space is related to the sameness of investment actions between two mutual funds. Conversely, spatial nearness in an Euclidean dissimilarities solution space is related to the sameness of investment inactions between two mutual funds. Because zero entries prevail in the matrix of investment preference profiles, scaling solutions based upon an Euclidean proximity formula tend to be dominated by the effect of these negative features of fund portfolio profiles. Therefore, the low stress values associated with the Euclidean proximity solutions are specious readings, reflecting no more than the fact that a close-packed set of points is a good reproduction of the low dissimilarities between funds in terms of the investments that they do not make.¹

These arguments are compelling to accept holdings overlap and reject Euclidean distance as the appropriate formula for computing

1. The distinct horizontal trend in the Euclidean proximity solution space can be traced solely to the exaggerated (squared) effect of the long-term debt asset group which tends to be under-represented in strong growth-oriented funds and highly represented in income-oriented funds.

pairwise proximities between portfolio investment profiles. In the words of Myron Wish, "A good fit to the 'wrong' data may be worse than a poor fit to the 'right' data." [7]

Beyond empirical arguments, an axiomatic foundation also is available to support the use of holdings overlap. Restle [8], Gregson [5] and others have developed overlap-type similarities from a set theoretic approach. In this sense, similarity evaluation becomes a matching process related to the intersection of attribute sets for pairs of subjects. Although set intersections do not satisfy metric space requirements (triangle inequality) except perhaps locally, Restle has indicated how adequate distance measures may be derived from set theoretic measures. Investigations have shown that scaling procedures are robust when distance functions fail even to obey the triangle inequality axiom [9]. Further discussion of these points is beyond the intent of the present study. For more information the reader is referred to the cited references.

NON-EUCLIDEAN SOLUTION SPACES

The previous topic may be regarded as a quest for an appropriate formula to compute distances between mutual fund subject points in the original 97-dimensional investment group stimuli space. Once computed, these proximity values are used in a multidimensional scaling algorithm to configure the fund points in a reduced-dimensionality metric solution space, the axes of which correspond to the basic

"psychological" attribute factors of mutual fund evaluation. The present topic, related to the previous one, concerns the choice of appropriate distance function in the final solution space.

The notion of "metric" space is one in which a distance function is defined, that function obeying four specific axiomatic conditions (the triangle inequality is one; see page 38. For a complete discussion of these metric space distance axioms, see Bechenbach and Bellman [10]). Euclidean distance, computed by the Pythagorean theorem, is but one such possible distance function. A particular class of distance functions is known as the Minkowski r -metric. This one-parameter class of distance functions is defined as follows:

$$d_r(x, y) = \left[\sum |x_i - y_i|^r \right]^{1/r}, \quad r \geq 1$$

where x and y are two points in n -space and $d_r(x, y)$ is their r -dependent distance. From this it can be seen that Euclidean distance is but the special case where $r = 2$.

Another interesting case is where $r = 1$ and is known as the city-block or Manhattan metric:¹

$$d_1(x, y) = \sum |x_i - y_i|$$

1. Refer to formula 2 on page 111.

With this formula, distance between two points depends only on the absolute differences in their coordinates, dimension by dimension. It gets its name because to get between two points in a city, one must travel along perpendicular lines parallel to the street and avenue "coordinates"; diagonal movements are not possible.

Specifying which distance function is appropriate for a particular multi-dimensional solution space requires an understanding of the underlying psychological processes of perception and evaluation. Torgerson has commented on the city-block metric as follows:

"... If a subject is required to rate a set of stimulus pairs with respect to their similarity, and the stimuli differ with respect to obvious and compelling dimensions, his ratings might very well behave as though they were a straight sum of the differences on the separate dimensions: that is, as though the subject were saying, 'These two stimuli differ this much with respect to brightness, plus this much with respect to size, plus this much with respect to shape.'

"On the other hand, if separate dimensions are not obvious, the subject might be more likely to judge the overall difference directly. In this event we would expect the Euclidean model to show more promise." [11]

In most instances, the psychological process underlying stimuli similarities judgments is only conjecture and, in the case of derived proximities from quasi-behavioral profile data as in the present study, it is difficult to predict the effect of the data transformation step. Consequently, selecting an appropriate metric for a solution space usually can proceed only on a trial and comparison basis in a search for improvement in solution goodness-of-fit and interpretability. The M-D-SCAL

computer program offers the convenient feature of allowing the analyst to specify Minkowski space solutions for alternative r values.

Scaling experiments were repeated to see if better stress and more interpretable solutions could be found than that of the Euclidean space experiment of Chapter 3. City-block scaling solutions were explored over a range of dimensionalities using proximities based upon First Quarter 1971 portfolio holdings data. Mutual fund proximities were calculated both as holdings overlap similarities and city-block dissimilarities (see the previous section on alternative proximity measures, page 110). Following are the resulting solution space stress results to be compared against the standard Euclidean space solutions in Table 8, page 112:

Table 9

Stress of Scaling Solutions Within City-Block
Metric Space for Alternative Proximity Measures

<u>Dimensionality</u>	<u>Proximity Measures</u>	
	<u>Holdings Overlap Similarities</u>	<u>City Block Dissimilarities</u>
1	0.514	0.515
2	0.416	0.414
3	0.338	0.346
4	0.300	0.305

For both proximity measures over all dimensionalities the Euclidean distance metric (Table 8) gave better fitting configurations. Visually, the city-block metric configurations represented only slight distortions

of the Euclidean metric displays. Thus, of the two Minkowski metrics investigated, r equals 1 or 2, the Euclidean space results were judged to be superior.

JOINT SPACE MODELS--PREFERENCE SCALING

Each mutual fund portfolio is an amalgam investment of various industry and asset group holdings. As stated in the Introduction, every fund portfolio can be thought of as a point positioned in investment group perceptual space in such a way that a fund's distance from an industry group point is proportional to the magnitude of the fund's holdings in that group. Suggested is the notion of a joint display of mutual fund subjects and industry and asset group stimuli in the same reduced factor space.

A large amount of interesting information would be obtained if we could unravel the joint subject-stimuli space of mutual fund holdings. Scale values for the Income and Growth attributes of individual industry and asset group stimuli would be produced as they are for the mutual fund subjects.

Various methods are available to achieve a display of subjects and stimuli as points or vectors in a common factor space. They all depend one way or another, on preference ratings for test stimuli on an individual or paired comparison basis as expressed by a panel of subjects. Thus, joint space models are known as "preference scaling" techniques.

Preference scaling and similarities (proximities) scaling constitute the major dichotomous classification of multidimensional scaling techniques generally.

Carroll [12] distinguishes between internal and external modes of preference scaling. Internal methods obtain results from a direct decomposition of the information contained in a stimuli-by-subjects preferences matrix like our Table A3 of fund portfolio profiles. A joint space solution of stimuli and subjects is produced simultaneously. Discussion of the application of these internal preference scaling techniques in the present study is delayed until page 127.

An external analysis of preference starts with an a priori subjects space, say, as output from a M-D-SCAL experiment on a subjects-by-subjects similarities matrix. Independently, each subject's preference ratings are determined for all of a set of test stimuli. These data take the form of a rectangular stimuli-by-subjects preferences matrix. Each stimulus row in the matrix of preferences can be thought of as a stimulus property vector across the sample of subjects. Thus, in a manner conceptually and procedurally identical to fitting property vectors (see page 80) each stimulus can be mapped onto the a priori subject space in a best fit fashion.¹

1. Insofar as the mathematics of any internal or external joint space model is concerned, the designation of the rows of the preference matrix as subjects or stimuli (and conversely, the columns as stimuli or subjects) is entirely arbitrary.

Joint Space Experiment
Mixed Preference Scaling

A method for generating a joint mutual fund-industry group space has been designed for the present study which combines both external and internal features. The 97 stimuli groups have been mapped as vectors onto the mutual fund subject space using the rows in Table A3 as property or preference scale values. The subject space is our standard M-D-SCAL two-dimensional solution, Figure 2 in Chapter 3, derived from holdings overlap proximities.

This procedure is external in the sense that subject points and stimuli vectors are not produced simultaneously but that the stimuli vectors are mapped onto an a priori subject space. It is internal in the sense that the matrix of preference values for mapping the stimuli is the same data (Table A3) from which the holdings overlap similarities were derived.

Mapping the stimuli vectors in two-dimensional fund space was accomplished using the linear multiple regression option of the PROFIT computer program. In accordance with the argument that a zero holding in an industry group by a fund is an unreliable measure of the fund's preference for that group, the PROFIT program was used in a way to disregard all zero holding data values.

Table A11 presents the results of the PROFIT analysis. For each of the 97 investment groups it shows: the number of funds out of 46 that

held securities in the group; the two dimensional direction cosines for each fitted investment group with respect to the horizontal (X) and vertical (Y) axes in the original M-D-SCAL solution space; and the correlation coefficient between the fitted and original vectors. Figure 15 is a plot of the joint space showing only some of the investment group vectors, the ones whose positions are most reliable statistically.¹ Of particular interest is the vector for Group 96 representing long term debt securities. At a positive angle of 16 degrees with respect to the X axis, it is quite collinear with the yield property vector (29 degrees, see page 83) and, as such, provides additional confirmation of an "Income" label for the abscissa in the rotated M-D-SCAL solution display.

The large number of fund vectors absent from Figure 15 reflects the fact that for most industry groups only a minority of funds have non-zero holdings. The average group is held by 17 of the 46 subject mutual funds. Some industry groups are represented in as few as three or four of the 46 subject portfolios. Because of this, the industry group stimuli space must be regarded as being much less reliable statistically than the subject mutual fund space. In the M-D-SCAL fund space each subject fund point has been positioned by 45 data values representing the interpoint proximities to the other funds. Good statistical reliability for all the industry group vectors could only be achieved by a considerable enlargement of the sample beyond 46 subjects.

1. A vector was included in the plot only if the hypothesis that its true correlation coefficient is zero could be rejected at the 99% confidence level.

Figure 15

INVESTMENT GROUP VECTORS
FITTED WITHIN M-D-SCAL MUTUAL FUND SPACE

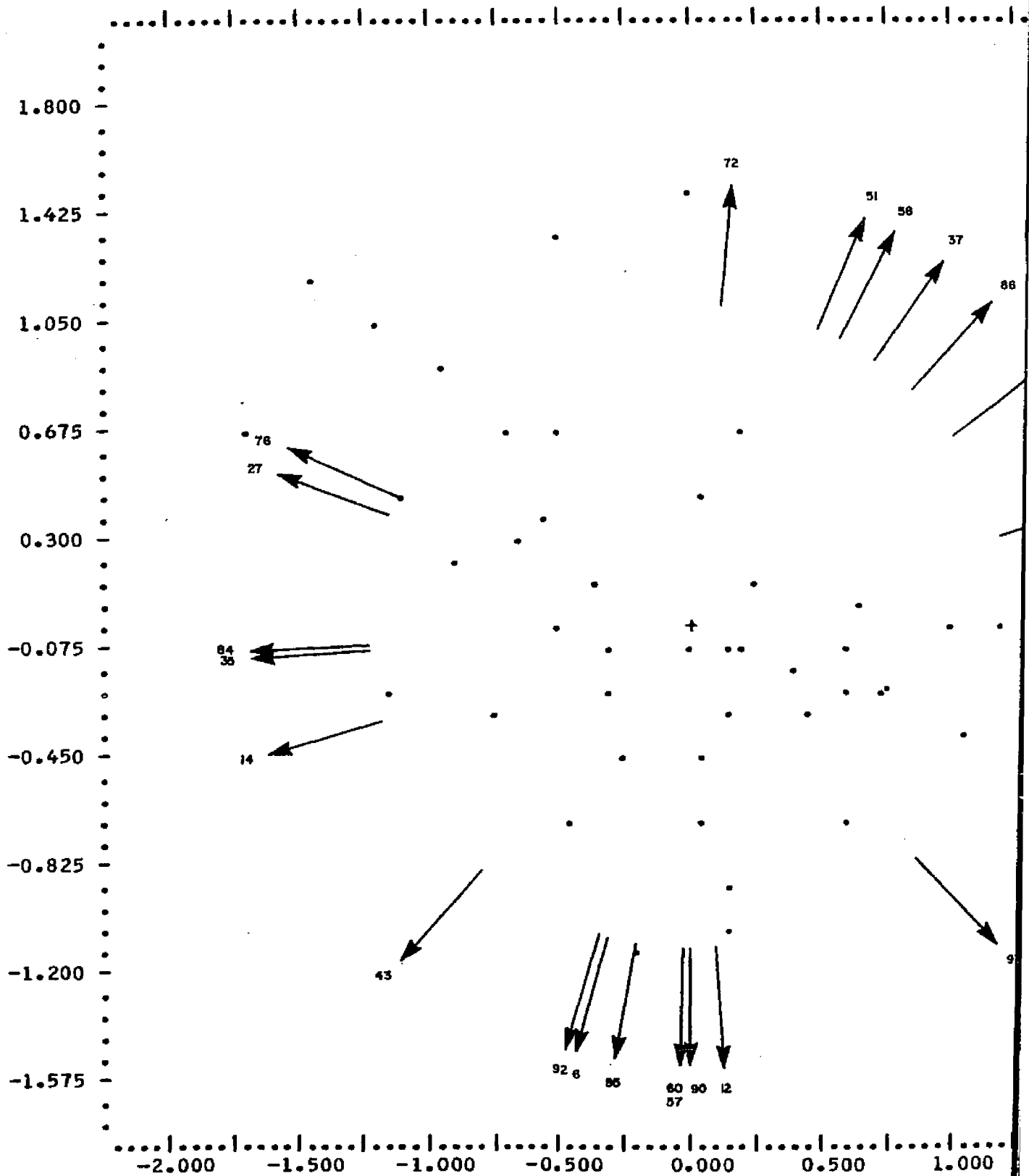
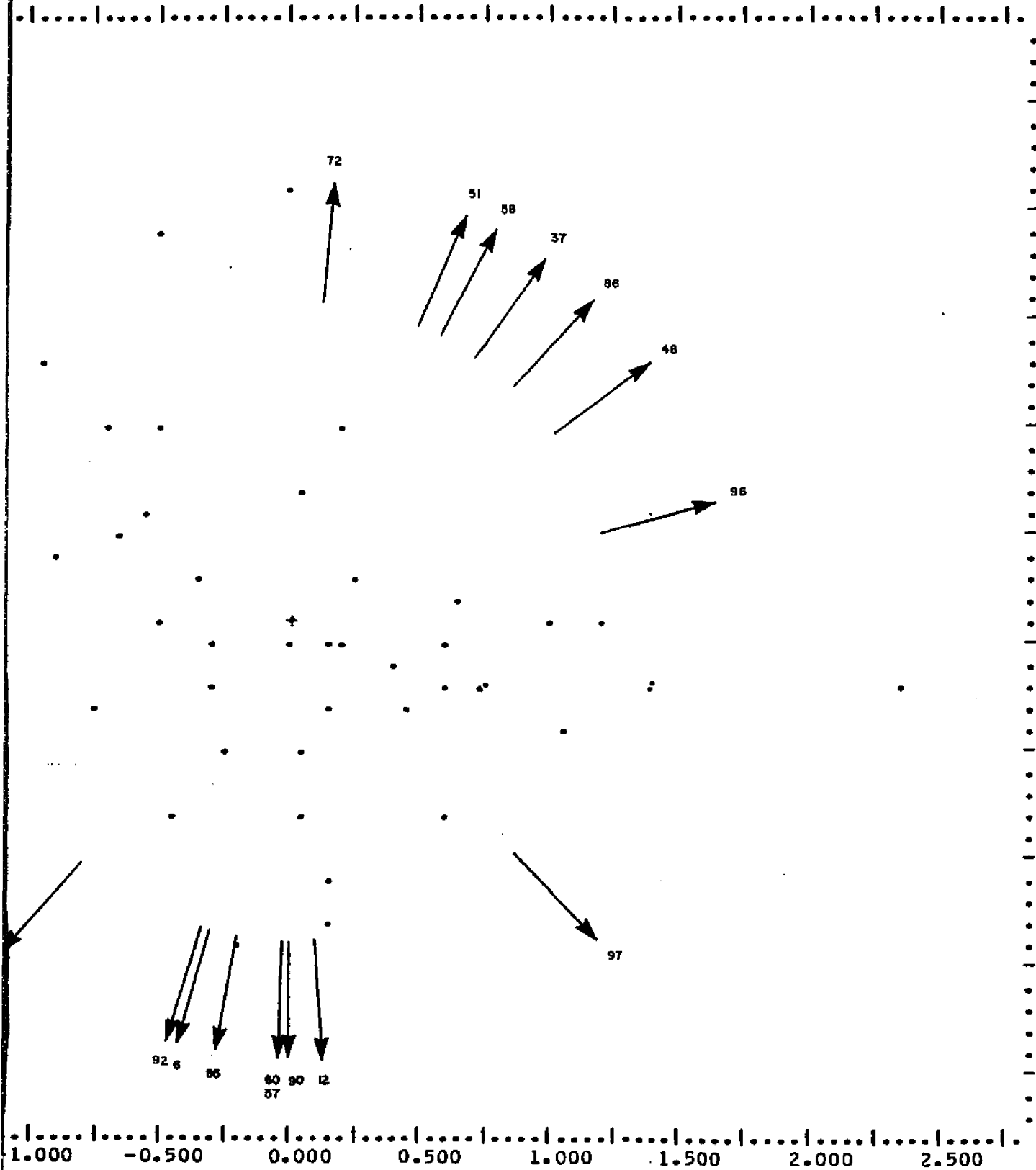


Figure 15

INVESTMENT GROUP VECTORS
FITTED WITHIN M-D-SCAL MUTUAL FUND SPACE



Joint Space Experiment
Internal Preference Scaling

Internal preference scaling methods mentioned earlier are capable of producing joint subject-stimuli spaces directly from a stimuli-by-subjects preferences matrix. Several internal preference scaling models are available embodied in different computer programs.

A major distinction among internal preference models are: 1) those that represent both subjects and stimuli as points in the common factor space and, 2) those that represent subjects as points and stimuli as vectors. The first case is known as the "ideal point" or "unfolding" model. Computer programs based on this model find an ideal point for a subject such that the distances from the point to the set of stimuli points represents a proportional best fit to the subject's input preference ratings for the stimuli. The subject's ideal point can be regarded as a hypothetical stimulus having the combination of attributes which are ideal for the subject. An example of this model is the CANDECOMP option of the INDSCAL computer program discussed at the end of Chapter 4. CANDECOMP is not limited two-variable joint space analyses, but is capable of handling multi-way input data to produce up to seven-variable joint space solutions (Chapter 4 is an example of its use in a three-way case).

The second of the two internal preference scaling models is called the vector model. A computer program based upon this model finds a set of stimuli points and a vector direction for each subject in the same

space such that the projections of the stimuli on a vector best agree with the preference ratings of the stimuli by the subject.

Actually, the ideal point and vector models are not to be contrasted. The vector model can be viewed as a special case of the ideal point model where the subjects' ideal points are assumed to extend to infinity.

The joint space capability of these internal preference scaling methods fits our research objective for finding a mutual fund-industry and asset group common factor solution. Our rectangular stimuli-by-subjects preferences matrix is in the correct form for input to these programs. Accordingly, an internal preference scaling experiment design has been explored in the present study.

Carroll and Chang's MDPREF [13] vector model computer program was used. MDPREF is based upon an Eckart-Young decomposition procedure (a type of principle components factor analysis). The First Quarter 1971 matrix of investment group preferences (Table A3) was submitted as input for an analysis in two dimensions. The resulting coordinates for the 97 investment group stimuli points are given in Table A12(a) and the direction cosines for the 46 mutual fund subject vectors are given in Table A12(b).

Figure 16 is a plot of the joint space investment groups points and fund vectors as it is produced by MDPREF (the reverse of the representation in the joint space display of Figure 15). Most stimuli points

Figure 16

MDPREF JOINT SPACE SOLUTION
SHOWING INVESTMENT GROUP POINTS AND MUTUAL FUND VECT

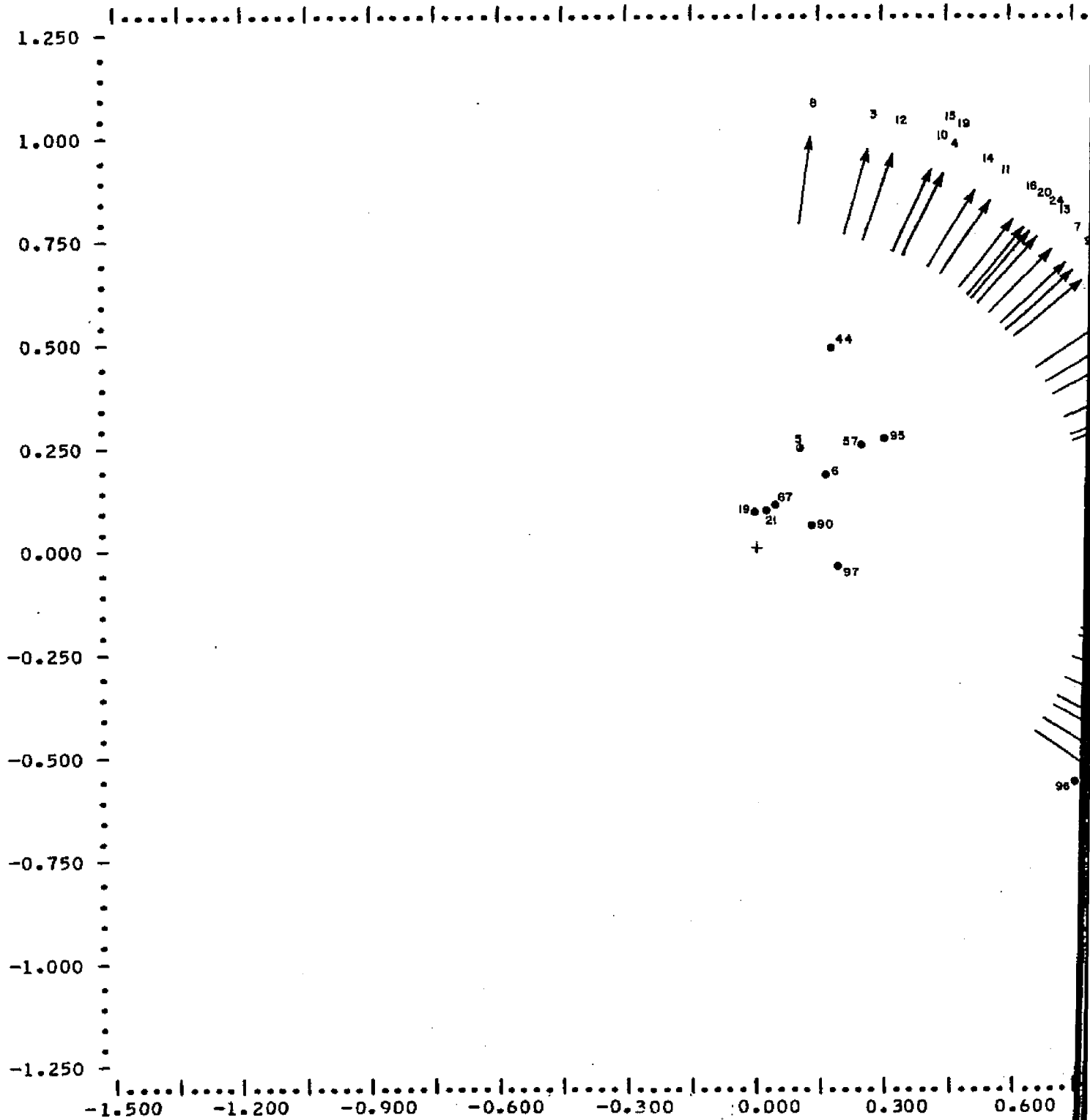
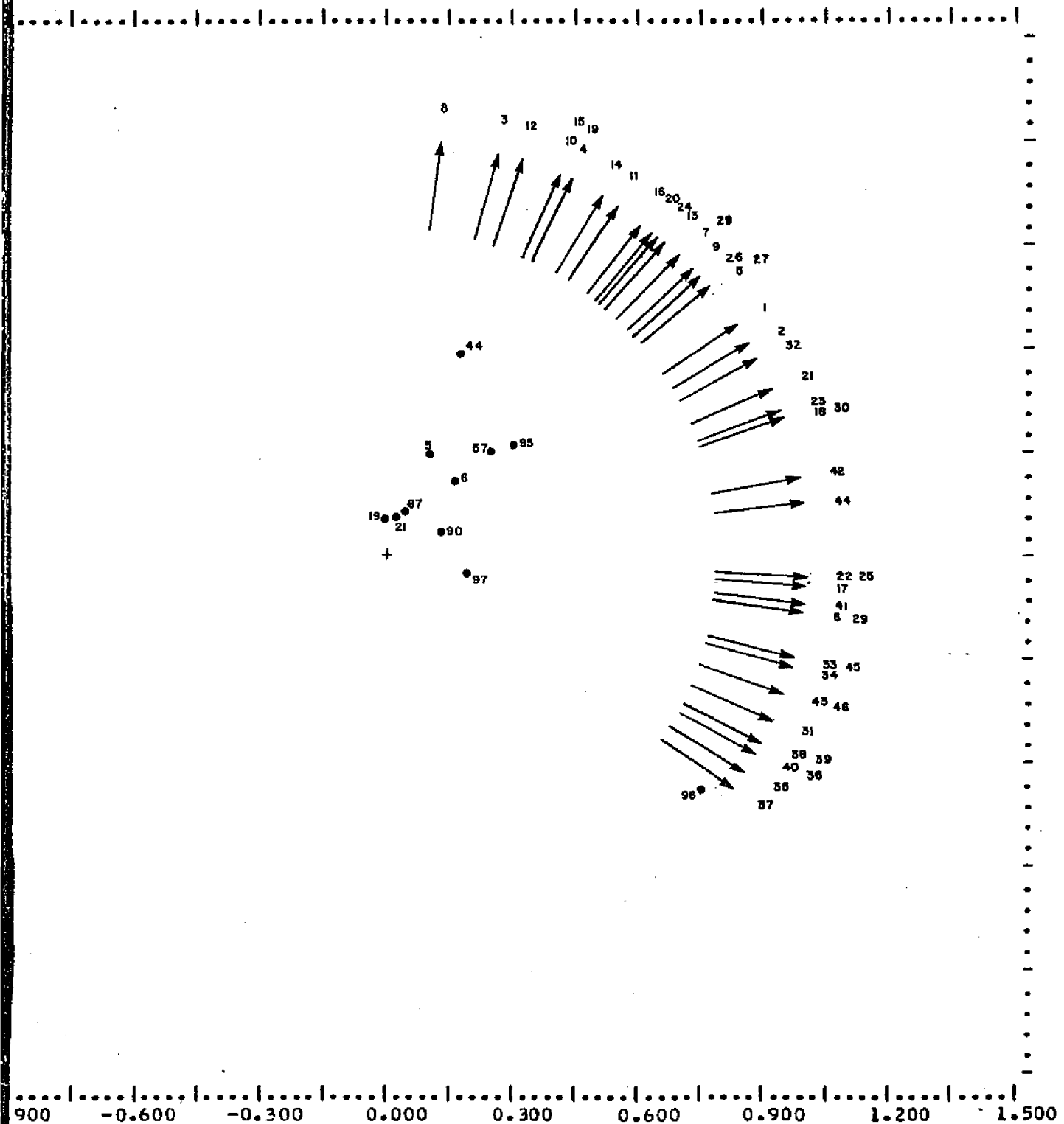


Figure 16

MDPREF JOINT SPACE SOLUTION

OWING INVESTMENT GROUP POINTS AND MUTUAL FUND VECTORS



clustered indistinguishably in the center of the display not showing any significant weighting on either dimension. Only those few industry and asset groups having meaningful dimension weights are plotted (note particularly the long term debt category, stimuli point 96).

Even without rotation, the axes in the MDPREF solution show a noticeable correspondence to the earlier M-D-SCAL results, namely, Income as the abscissa and Growth as the ordinate. This can be seen in the tendency of the subject vectors to describe an arc ranging from G_1 and G_2 class funds (numbers 1 to 16) oriented vertically to I class funds (numbers 33 to 40) oriented horizontally. A clockwise orthogonal rotation of the axis system through about 25 degrees seems appropriate for improved interpretability.

That the rank order of mutual fund scale values on Growth and Income dimensions from the MDPREF solution shows many differences with their rank order in the M-D-SCAL solution of Figure 2 is to be expected. The mathematical procedures underlying the internal preference scaling model and the similarities scaling model are different. This means that the M-D-SCAL results cannot be accepted as correct without an explanation of the discrepant MDPREF results.

An answer lies in the fundamental assumptions behind internal preference scaling models concerning the nature of the input preference data. As Coombs [14] has pointed out, these models assume that all test stimuli are potential first choices. That is, the rank order for a stimulus

is obtained as the expressed preference for the item given that all other stimuli are not available. Preference ratings of this sort fall within Coomb's concept of "dominance" data.

Careful consideration of the holdings percentage data shows them not to be first choice preferences. Instead, the percentage of an investor's total assets that he invests in the securities of a particular industry group is thought to be influenced by the total package of his investment selections. As explained in the Introduction, this does not necessarily mean that a bundle of securities is selected taking their mutual interrelations into consideration. More simply, it could mean that marginal preference for a single investment decreases as the holding becomes a larger and larger percentage of the investor's total assets. (This is the model discussed on page 12.)

No matter what the exact process of portfolio selection, it is clear that our data in Table A3 are not first choice preferences in the sense described by Coombs. How then should this data be regarded? Shepard [15] has put forth a taxonomy of multidimensional scaling methods based upon four basic types of data. In addition to proximity data and conjoint data,¹ he distinguishes between dominance data and profile (multivariate) data. Everything considered, it is most appropriate to regard portfolio holdings

1. Conjoint data: a rectangular matrix where "each entry is the magnitude of an effect that arises jointly when the two contributing variables take on the levels corresponding to the row and column of that entry." [16]

as profile data across which mutual fund proximities can be computed. The transformation into proximities must be based upon a formula, holdings overlap, which reconciles with the process of investment budgeting. From this, the data preparation steps of Chapter 2 and experiment design of Chapter 3 logically follow.

REFERENCES
Chapter 5

1. L. C. Cole, "The Measurement of Partial Interspecific Association," Ecology, Vol. XXXVIII (1957), pp. 226-233.
2. C. H. Coombs, A Theory of Data (New York: John Wiley & Sons, Inc., 1964).
3. P. Green and V. Rao, "A Note on Proximity Measures and Cluster Analysis," Journal of Marketing Research, Vol. VI (August, 1969), pp. 359-364.
4. R. Sokal and P. Sneath, Principles of Numerical Taxonomy, (San Francisco: W. H. Freeman and Company, 1963).
5. R. A. M. Gregson, "Quadratic Similarities," British Journal of Mathematical and Statistical Psychology, Vol. XXIII (May, 1970), pp. 53-68.
6. R. Sokal and P. Sneath, Principles of Numerical Taxonomy, pp. 128-130.
7. Myron Wish, "Notes on the Variety, Appropriateness, and Choice of Proximity Measures," mimeographed paper prepared for the Bell-Penn Workshop on Multidimensional Scaling (Philadelphia: University of Pennsylvania, June 7-10, 1972), p. 22.
8. F. Restle, Psychology of Judgment and Choice, (New York: John Wiley & Sons, Inc., 1961).
9. P. E. Green and F. J. Carmone, Multidimensional Scaling and Related Techniques in Marketing Analysis, Marketing Science Institute Series (Boston: Allyn and Bacon, Inc., 1970), p. 127.
10. E. Beckenbach and R. Bellman, An Introduction to Inequalities (New York: Random House, 1961).
11. W. S. Torgerson, Theory and Methods of Scaling (New York: John Wiley & Sons, Inc., 1960), p. 258.
12. J. Douglas Carroll, "Individual Differences and Multidimensional Scaling," in Multidimensional Scaling: Theory and Applications in the Behavioral Sciences, Volume I, Theory, ed. R. N. Shepard, et al (New York: Seminar Press, 1972).

13. J. J. Chang and J. D. Carroll, "How to Use MDPREF, a Computer Program for Multidimensional Analysis of Preference Data," mimeographed report (Bell Telephone Laboratories, Murray Hill, N. J., 1969).
14. C. H. Coombs, A Theory of Data (New York: John Wiley & Sons, Inc., 1964).
15. Roger N. Shepard, "A Taxonomy of Some Multidimensional Scaling Methods," mimeographed paper prepared for the Bell-Penn Workshop on Multidimensional Scaling (Philadelphia: University of Pennsylvania, June 7-10, 1972).
16. Ibid., p. 4.

Chapter 6

SUMMARY

This report began with an explanation of the need to provide portfolio investors a new and better way of rating the investment attributes of existing mutual funds. Discovering an improved approach was shown to be equivalent to answering such fundamental questions as "What are the true perceptual dimensions of investment preference?" and "How can we scale existing portfolios with respect to these preference dimensions once found?" Although being knotty problems requiring "objective" measurements of ex ante variables, the alternative of avoiding these questions is unsatisfactory. Despite the apparent difficulties, this report reflects an attempt to come to grips with the problem and find a workable empirical solution procedure. The following states in simple terms what has been accomplished.

A procedure has been presented which gives a continuous plot of mutual funds within a coordinate system that represents the basic dimensions of investment preference. Remarkably, this spatial display is produced using only the information contained in a table of the funds' portfolio holdings, the holdings aggregated into standard industry and asset groups. No external information is necessary as to past performance or predicted future performance of the funds or their component investments. The procedure is "ignorant" as to the identity of the standard investment groups; indeed, it only uses single-value fund similarities

computed across the profiles of fund holdings, and then only the rank order relationships in the similarities data are used, not the numbers themselves.

Cluster patterns in the spatial display of mutual funds and property fitting tests have shown persuasively that the dominant dimensions of investment preference correspond to the conventional notions of income and growth investment objectives. Most important, scale values on the Growth dimension truly represent current growth expectations, not surrogates for expectations of growth. They are expectations in a consensual sense, distilled from the pattern of investment concentrations from a cross-section of mutual fund portfolios representing a diversity of investment objectives.

In sum, the nonmetric multidimensional scaling-based procedures take weak, readily available, "objective" portfolio similarities data and upgrade it to reveal metric-scaled fund investment attributes--stronger and more meaningful information than is produced by any other method of evaluating portfolios. With this information, an investor can identify the mutual fund whose portfolio best meets his combination of investment objectives. The display of information can be updated as readily as new data becomes available concerning portfolio changes.

Perhaps of equal significance to the mutual fund plot is the ability to portray individual industry and asset groups within the revealed investment objective coordinate system. The investment group space

is produced from the same data which is used to develop the fund space. Although of less statistical reliability than the mutual fund results, the display of investment groups in the common evaluative space can provide portfolio architects--fund managers and individual investors--vital information concerning currently perceived investment attributes of individual industry groups.

A third potential application of the study results is a method to characterize the investment attributes of the overall market at a given point in time. To the extent that the consensus change in investment objectives for a large sample of mutual funds mirrors the change in orientation for the market as a whole, a new barometer of market attitudes is suggested using the configuration matching procedure for comparing scaling solutions over time.

In light of these results, the companion objectives of this study can be said to have been achieved:

- 1) To develop a method for empirically displaying the factor structure of portfolio differences and,
- 2) To explore the performance of multidimensional scaling tools under somewhat new conditions and data.

For both of these points the present study is no more than just a beginning. Despite the large amount of research that was conducted, by no means has this investigation exhausted the experimental variations that reasonably can be considered nor has it fully probed the implications of

the variations that were tested. Beyond the descriptive and classificatory purposes of the present study, the opportunity has been opened to investigate how the revealed factors explain or cause investor portfolio selection behavior. Among the related questions remaining to be answered is an interpretation of higher level factors. No connection has been found between any feature of the empirical factor structure and formal notions of risk. The empirical evidence and the force of logic tends to support the belief that risk does not exist as a separate salient factor in the perspective of portfolio differences. However, whether indeed risk does not actually exist as a separate dimension of perception; whether separate aspects of risk are confounded with the Growth and Income dimensions; whether risk is perceived as a region in Growth/Income space; or whether the experimental design is simply not capable of detecting the risk effect are questions remaining to be settled. Hopefully, the present study will serve to prompt continued investigations to explore these and other questions in the area of finance using the promising techniques of multidimensional scaling.

APPENDIX

Table A1

EXPERIMENT SAMPLE OF 46 MUTUAL FUND SUBJECTS
Showing Their Investment Objective Classifications

		Investment Objective Classifications			
	Total Assets 3/31/71 (\$000,000)	Fund- Scope	Arthur Lipper	Standard & Poor's	Wiesen- berger
Fund					
G ₁ (Growth Funds):					
1 Essex	55.1	G-pf ¹	G	G	MCG ²
2 Fairfield	57.1	G-pf	G	G	MCG
3 Channing Special	113.3	G-pf	G	G	MCG
4 Colonial Equities	136.8	G-pf	G	G	MCG
5 Scudder Special	209.8	G-pf	G	G	MCG
6 Winfield Growth	147.3	G-pf	G	G	MCG
7 American Investors	250.5	G-pf	G	G	MCG
8 Keystone S-4	698.5	G-pf	G	G	MCG
G ₂ (Growth Funds):					
9 Capital Shares	53.6	G-dcs ³	G	G	G
10 Anchor Growth	512.5	G-dcs	G	G	G
11 T. R. Price Growth Stock	801.4	G-dcs	G	G	G
12 Keystone S-3	186.6	G-dcs	G	G	G
13 Channing Growth	274.4	G-dcs	G	G	G
14 Fidelity Trend	1124.9	G-dcs	G	G	G
15 National Investors	851.3	G-dcs	G	G	G
16 Egret Growth	38.2	G-dcs	G	G	G
GI (Growth-Income Funds):					
17 Pioneer	214.0	GI	GI	GI	GI
18 Guardian Mutual	59.6	GI	GI	GI	GI
19 Keystone S-1	56.4	GI	GI	GI	GI
20 One William Street	278.9	GI	GI	GI	GI
21 Hamilton H-DA	693.7	GI	GI	GI	GI
22 Fidelity	898.0	GI	GI	GI	GI
23 Affiliated	1736.1	GI	GI	GI	GI
24 Massachusetts Investors Trust	2204.8	GI	GI	GI	GI

(Continued)

1. "Performance" Growth Funds
2. "Maximum Capital Gains" Growth Funds
3. "Diversified Common Stock" Growth Funds

Table A1 (Continued)

		Total Assets 3/31/71 (\$000,000)	Investment Objective Classifications			
Fund	Fund- Scope		Arthur Lipper	Standard & Poor's	Wiesen- berger	
IG (Income-Growth Funds):						
25	Everest	53.5	IG	GI ⁴	IG	IG
26	Broad Street Investing	408.2	IG	GI	IG	IG
27	Washington Mutual Investors	324.4	IG	GI	IG	IG
28	Keystone S-2	134.4	IG	GI	IG	IG
29	Colonial	240.3	IG	GI	IG	IG
30	Mutual Investing Foundation (MIF)	202.5	IG	GI	IG	IG
31	Financial Industrial Income	114.2	IG	GI	IG	IG
32	Southwestern Investors	28.6	IG	GI	IG	IG
I (Income Funds):						
33	Puritan	877.5	I	I	I	I
34	American Express Income	149.0	I	I	I	I
35	Keystone K-1	109.7	I	I	I	I
36	Value Line Income	93.9	I	I	I	I
37	Northeast Investors	45.0	I	I	I	I
38	Decatur Income	145.8	I	I	IS	I
39	Provident Fund for Income	110.5	I	I	ISG	I
40	National Securities- Income	87.9	I	I	ISG	I
GIS (Growth-Income-Stability Funds):						
41	Loomis-Sayles Mutual	170.5	GIS	B ⁵	GIS	GIS
42	Stein Roe & Farnham Balanced	171.7	GIS	B	GI	GIS
43	Channing Balanced	102.2	GIS	B	GIS	GIS
SIG (Stability-Income-Growth):						
44	Wellington	1447.5	SIG	B ⁵	SIG	SIG
45	Massachusetts	217.3	SIG	B	SIG	SIG
46	Nation-Wide Securities	85.6	SIG	B	S	SIG

4. Arthur Lipper's Classification System Does Not Include an IG Category

5. Balanced Fund

Table A2

STANDARD INVESTMENT GROUPS

1. Aircraft & Parts
2. Apparel & Other Textiles
3. Banking & Brokers
4. Cement - Hydraulic
5. Chemicals - Drugs
6. Chemicals - Industrial, Plastics
7. Chemicals - Soaps, Toilet
8. Chemicals - Other
9. Coal Mining
10. Communications - Telephone, Telegraph
11. Communications - Radio, Television
12. Concrete, Gypsum & Plaster
13. Conglomerates
14. Contract Construction
15. Credit - Personal & Business
16. Credit - Savings & Loan
17. Electrical - CATV
18. Electrical - Communications
19. Electrical - Electronics
20. Electrical - Household
21. Electrical - Industrial, Testing
22. Electrical - Radio, T. V., Other
23. Fabricated Metal Products
24. Food Products - Bakery, Confectionary
25. Food Products - Beverages
26. Food Products - Canned, Frozen
27. Food Products - Dairy
28. Food Products - Grain Mill
29. Food Products - Meat
30. Food Products - Sugar
31. Furniture & Fixtures
32. Glass & Glassware
33. Hotels & Other Lodging
34. Instruments - Scientific, Watches & Clocks
35. Instruments - Optical & Photographic
36. Insurance Carriers & Agents
37. Investment Companies
38. Investment Institutions - Other
39. Leather & Leather Products
40. Lumber & Wood Products
41. Machinery - Farm, Construction
42. Machinery - General Industrial
43. Machinery - Metal Working
44. Machinery - Office & Computer
45. Machinery - Service Industry
46. Machinery - Special Industry
47. Manufacturing Industry - Other
48. Metal Mining - Copper, Lead, Zinc & Other

Table A2 (Continued)

49. Metal Mining - Gold
50. Minerals - Nonmetallic
51. Motion Pictures & Other Recreation
52. Motor Vehicles
53. Motor Vehicle Equipment
54. Oil & Gas Extraction
55. Oil & Gas Field Services
56. Paper & Allied Products
57. Petroleum Refining
58. Pollution Control
59. Primary Metals - Steel & Iron
60. Primary Nonferrous - Aluminum
61. Primary Nonferrous - Copper
62. Primary Nonferrous - Other
63. Printing & Publishing - Books, Periodicals
64. Printing & Publishing - Other
65. Real Estate
66. Retail - Apparel & Accessories
67. Retail - Department Stores, Mail Order
68. Retail - Drug Stores
69. Retail - Eating & Drinking
70. Retail - Food Stores
71. Retail - Variety Stores
72. Retail - Other
73. Rubber - Tires
74. Rubber & Plastics - Fabricated, Other
75. Services - Automobile
76. Services - Computer Leasing, Software
77. Services - Nursing Homes
78. Services - Personal, Other
79. Stone, Clay, Glass - Other
80. Textile Mill Products
81. Tobacco Manufacturing
82. Toys
83. Trailer Coaches - Mobile
84. Transportation - Air
85. Transportation - Railroad
86. Transportation - Other
87. Transportation Equipment - Railroad
88. Transportation Equipment - Other
89. Trucking & Warehousing
90. Utilities - Electric
91. Utilities - Gas
92. Utilities - Combination, Water
93. Wholesale - Electrical
94. Wholesale - Chemical & Other
95. Cash & Near Cash
96. Long-Term Debt including Convertibles
97. Preferred Stock including Convertibles

MUTUAL FUND PORTFOLIO HOLDINGS
PERCENT OF TOTAL ASSETS

GROUP	FUND																																													
	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46																							
1	.006	.056	.000	.025	.000	.000	.000	.000	.043	.006	.026	.000	.000	.015	.024	.000	.000	.003	.000	.002	.000	.000																								
2	.000	.000	.000	.000	.000	.007	.030	.010	.000	.000	.016	.000	.016	.000	.000	.061	.000	.007	.000	.000	.000	.007	.007																							
3	.036	.000	.000	.000	.085	.032	.000	.013	.000	.013	.000	.019	.000	.011	.008	.000	.032	.008	.047	.035	.060	.006	.020																							
4	.004	.000	.000	.013	.016	.000	.000	.013	.000	.003	.000	.000	.000	.000	.000	.009	.000	.008	.000	.000	.000	.000	.000																							
5	.042	.000	.097	.010	.033	.009	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.066	.058	.027	.019	.094	.000																							
6	.033	.047	.084	.130	.032	.073	.105	.054	.000	.053	.006	.000	.026	.000	.029	.020	.000	.039	.034	.000	.027	.042	.037																							
7	.011	.000	.015	.000	.028	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.019	.000	.000	.026	.000	.014																							
8	.000	.000	.000	.000	.000	.012	.014	.000	.000	.000	.000	.000	.000	.000	.000	.005	.000	.027	.011	.000	.000	.000	.000																							
9	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000																							
10	.043	.020	.000	.040	.004	.021	.026	.036	.000	.028	.031	.000	.006	.000	.018	.015	.000	.017	.000	.020	.038	.016	.013																							
11	.014	.000	.000	.019	.000	.009	.000	.000	.000	.008	.000	.000	.000	.000	.000	.000	.000	.000	.033	.000	.000	.000	.000																							
12	.016	.000	.000	.029	.019	.014	.022	.009	.000	.000	.000	.000	.000	.000	.000	.018	.000	.035	.000	.000	.000	.000	.000																							
13	.008	.000	.000	.000	.013	.015	.000	.000	.000	.000	.000	.000	.000	.000	.000	.009	.000	.000	.000	.000	.010	.000	.000																							
14	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.012	.000	.000	.000	.000	.008	.000	.000	.026	.000	.000	.000	.000																							
15	.017	.000	.020	.012	.000	.034	.000	.007	.029	.004	.028	.003	.012	.000	.000	.000	.024	.035	.000	.000	.009	.000	.009																							
16	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.012	.000	.006	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000																							
17	.002	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000																							
18	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000																							
19	.005	.000	.014	.000	.000	.000	.000	.000	.000	.000	.000	.000	.004	.000	.000	.006	.000	.000	.000	.000	.013	.000	.000																							
20	.004	.000	.036	.000	.038	.018	.000	.014	.000	.010	.000	.000	.000	.000	.000	.000	.000	.009	.031	.000	.000	.000	.004																							
21	.000	.029	.076	.020	.009	.000	.019	.006	.000	.006	.004	.000	.000	.000	.000	.013	.000	.000	.031	.029	.012	.000	.007																							
22	.016	.000	.000	.004	.000	.019	.026	.000	.033	.000	.000	.000	.000	.000	.000	.000	.000	.008	.006	.000	.000	.012	.000																							
23	.010	.000	.014	.009	.000	.000	.000	.000	.000	.011	.009	.000	.000	.000	.000	.006	.000	.000	.017	.038	.000	.000	.011																							
24	.003	.000	.000	.000	.000	.000	.000	.000	.000	.010	.000	.000	.000	.000	.000	.004	.000	.000	.000	.000	.000	.000	.000																							
25	.035	.000	.011	.000	.000	.000	.000	.007	.000	.009	.000	.000	.013	.000	.000	.000	.000	.013	.000	.000	.000	.005	.000																							
26	.008	.000	.009	.009	.013	.006	.000	.007	.000	.000	.000	.000	.000	.000	.000	.008	.000	.012	.000	.020	.047	.000	.025																							
27	.000	.026	.000	.029	.000	.012	.006	.007	.000	.014	.006	.000	.000	.000	.000	.000	.000	.000	.000	.000	.019	.000	.013																							
28	.000	.000	.011	.000	.009	.000	.007	.000	.047	.014	.000	.000	.000	.000	.016	.000	.000	.000	.000	.000	.000	.000	.000																							
29	.005	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.008	.000	.000	.000	.000	.022	.000	.000	.000																							
30	.000	.000	.000	.000	.000	.000	.000	.000	.000	.002	.000	.000	.009	.000	.000	.000	.000	.000	.000	.017	.000	.000	.000																							
31	.003	.000	.000	.016	.000	.000	.000	.000	.000	.007	.010	.000	.009	.000	.000	.000	.000	.000	.027	.000	.000	.000	.000																							
32	.007	.000	.000	.000	.009	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.019	.000	.014	.000	.000	.011																							
33	.000	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.013	.000	.000	.000	.000																							
34	.000	.000	.000	.013	.008	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.011	.000	.000	.000	.000																							
35	.021	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.024	.012	.017	.000	.000	.000																							
36	.037	.000	.008	.000	.038	.000	.000	.008	.000	.005	.015	.019	.005	.000	.010	.000	.019	.030	.036	.043	.018	.000	.032																							
37	.006	.000	.000	.000	.000	.008	.000	.000	.000	.000	.000	.000	.000	.000	.000	.016	.017	.000	.000	.000	.005	.000	.000																							
38	.023	.012	.000	.000	.008	.023	.000	.009	.000	.005	.000	.027	.015	.000	.003	.000	.000	.010	.000	.000	.000	.003	.000																							
39	.000	.000	.018	.000	.000	.026	.000	.000	.000	.002	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.030																							
40	.017	.021	.009	.000	.000	.000	.000	.000	.053	.000	.000	.000	.000	.000	.000	.007	.000	.016	.054	.000	.000	.016	.000																							
41	.013	.000	.000	.012	.000	.006	.054	.000	.000	.019	.024	.009	.000	.000	.014	.000	.000	.000	.000	.015	.000	.000	.012																							
42	.000	.000	.013	.020	.000	.000	.000	.005	.000	.011	.008	.000	.008	.000	.000	.000	.000	.000	.000	.021	.000	.000	.000																							
43	.008	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.006	.000	.000	.000																							
44	.108	.025	.062	.000	.000	.052	.000	.000	.024	.000	.000	.000	.000	.000	.000	.011	.000	.061	.149	.000	.083	.048	.000																							
45	.004	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.008	.000	.000	.000	.000	.000	.000	.000																							
46	.034	.000	.005	.016	.000	.010	.000	.000	.031	.018	.008	.000	.005	.000	.000	.002	.020	.000	.000	.000	.005	.000	.000																							
47	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.007	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000																							
48	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000																							
49	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.008	.000	.000	.000	.000	.000	.000	.000																							
50	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000																							
51	.000	.000	.000	.000	.000	.000	.000	.000	.000	.010	.009	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000																							
52	.059	.000	.025	.023	.042	.000	.044	.034	.082	.022	.074	.049	.000	.000	.012	.000	.018	.029	.012	.033	.002	.000	.017																							
53	.000	.023	.019	.029	.017	.015	.081	.017	.035	.020	.009	.000	.000	.000	.039	.000	.017	.009	.000	.000	.000	.011	.000																							
54	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.008	.024	.000	.000	.000	.000	.024	.006	.000	.015	.000	.000																							
55	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.008	.000	.000	.000	.000	.000	.011	.000	.000	.000	.000																							
56	.010	.000	.020	.004	.000	.000	.060	.012	.013	.023	.000	.000	.006	.000	.009	.015	.000	.010	.000	.026	.000	.000	.039																							
57	.108	.049	.086	.049	.048	.083	.052	.075	.093	.062	.064	.039	.030	.000	.039	.021	.056	.086	.000	.050	.083	.055	.073																							
58	.030	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.011	.000	.000	.016	.000	.000	.009	.000																							
59	.034	.024	.000	.018	.000	.000	.013	.000	.000	.026	.019	.024	.000	.000	.013	.000	.023	.000	.000	.027	.000	.000	.006																							
60	.007	.000	.000	.022	.000	.000	.016	.000	.038	.000	.000	.000	.000	.000	.017	.000	.000	.000	.000	.000	.003	.000	.011																							
61	.009	.000	.000	.042	.000	.038	.016	.000	.000	.015	.000	.000	.000	.000	.004	.000	.000	.000	.000	.011	.010	.006	.000																							
62	.039	.000	.006	.000	.008	.035	.000	.000	.000	.005	.010	.000	.000	.000	.000	.010	.000	.013	.000	.004	.000	.003	.000																							
63	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.026	.0																																

Table A4

HOLDINGS OVERLAP SIMILARITIES BETWEEN MUTUAL

	Mutual Fund																								
Fund	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
2	0.22	0.32																							
3	0.23	0.32																							
4	0.29	0.46	0.30																						
5	0.29	0.31	0.18	0.30																					
6	0.24	0.34	0.28	0.31	0.32																				
7	0.23	0.24	0.19	0.31	0.27	0.32																			
8	0.21	0.42	0.44	0.45	0.31	0.40	0.33																		
9	0.14	0.28	0.22	0.29	0.29	0.34	0.23	0.22																	
10	0.20	0.29	0.38	0.46	0.41	0.46	0.40	0.50	0.35																
11	0.25	0.32	0.42	0.38	0.32	0.36	0.29	0.45	0.38	0.46															
12	0.28	0.31	0.34	0.48	0.43	0.41	0.34	0.44	0.38	0.53	0.38														
13	0.32	0.40	0.59	0.44	0.36	0.43	0.30	0.37	0.34	0.36	0.54	0.47	0.41	0.46											
14	0.32	0.23	0.29	0.36	0.43	0.30	0.37	0.34	0.36	0.47	0.46	0.38	0.42	0.52											
15	0.30	0.20	0.38	0.37	0.36	0.41	0.35	0.40	0.36	0.47	0.46	0.38	0.40	0.55	0.54										
16	0.23	0.28	0.36	0.37	0.37	0.36	0.35	0.37	0.41	0.53	0.45	0.39	0.40	0.55	0.54										
17	0.23	0.21	0.17	0.32	0.46	0.42	0.31	0.29	0.37	0.41	0.30	0.34	0.37	0.41	0.41	0.41									
18	0.25	0.40	0.21	0.32	0.50	0.42	0.37	0.35	0.38	0.43	0.40	0.36	0.43	0.46	0.34	0.43	0.49								
19	0.17	0.10	0.23	0.24	0.26	0.34	0.24	0.25	0.35	0.41	0.34	0.37	0.35	0.41	0.36	0.47	0.47	0.31	0.32						
20	0.29	0.28	0.33	0.39	0.36	0.40	0.43	0.37	0.35	0.50	0.54	0.42	0.47	0.58	0.56	0.53	0.37	0.44	0.44						
21	0.33	0.29	0.16	0.30	0.36	0.30	0.26	0.22	0.35	0.37	0.31	0.33	0.33	0.49	0.34	0.47	0.48	0.54	0.27	0.38					
22	0.25	0.22	0.24	0.32	0.46	0.44	0.25	0.23	0.39	0.46	0.33	0.41	0.44	0.44	0.34	0.51	0.51	0.49	0.45	0.38	0.47				
23	0.24	0.24	0.13	0.22	0.46	0.30	0.24	0.18	0.26	0.34	0.30	0.26	0.26	0.44	0.44	0.42	0.42	0.39	0.26	0.39	0.26	0.43			
24	0.27	0.26	0.14	0.25	0.43	0.31	0.29	0.18	0.34	0.39	0.54	0.45	0.38	0.47	0.57	0.47	0.41	0.44	0.56	0.51	0.51	0.60	0.51		
25	0.24	0.20	0.18	0.25	0.32	0.37	0.25	0.24	0.33	0.33	0.24	0.37	0.34	0.33	0.37	0.34	0.46	0.46	0.37	0.24	0.28	0.46	0.35	0.34	
26	0.22	0.21	0.21	0.30	0.41	0.37	0.30	0.25	0.38	0.43	0.39	0.30	0.39	0.47	0.58	0.54	0.36	0.42	0.42	0.48	0.45	0.46	0.52	0.35	
27	0.30	0.22	0.41	0.23	0.38	0.27	0.20	0.19	0.23	0.36	0.14	0.39	0.19	0.36	0.47	0.58	0.42	0.41	0.25	0.33	0.51	0.42	0.46	0.53	0.39
28	0.22	0.23	0.22	0.20	0.37	0.35	0.25	0.22	0.31	0.44	0.26	0.35	0.36	0.40	0.33	0.41	0.39	0.42	0.47	0.31	0.44	0.38	0.52	0.43	
29	0.18	0.20	0.21	0.20	0.34	0.36	0.26	0.22	0.21	0.37	0.26	0.29	0.35	0.40	0.32	0.48	0.44	0.51	0.25	0.33	0.52	0.54	0.42	0.51	0.49
30	0.19	0.26	0.10	0.22	0.47	0.41	0.21	0.13	0.33	0.25	0.25	0.22	0.34	0.24	0.34	0.43	0.53	0.22	0.38	0.65	0.34	0.63	0.45	0.32	
31	0.15	0.25	0.11	0.22	0.32	0.32	0.41	0.19	0.18	0.31	0.34	0.23	0.21	0.28	0.33	0.26	0.41	0.46	0.47	0.27	0.25	0.47	0.50	0.46	0.43
32	0.14	0.26	0.12	0.25	0.32	0.33	0.43	0.17	0.19	0.36	0.29	0.27	0.28	0.29	0.45	0.28	0.41	0.38	0.41	0.27	0.25	0.26	0.55	0.46	0.41
33	0.24	0.26	0.12	0.26	0.48	0.41	0.22	0.17	0.37	0.37	0.33	0.28	0.27	0.29	0.45	0.28	0.41	0.51	0.47	0.25	0.25	0.55	0.60	0.55	0.41
34	0.21	0.28	0.15	0.28	0.42	0.38	0.19	0.20	0.32	0.27	0.22	0.22	0.22	0.25	0.30	0.22	0.46	0.42	0.21	0.26	0.44	0.51	0.39	0.44	0.41
35	0.15	0.18	0.07	0.20	0.20	0.23	0.18	0.11	0.20	0.20	0.14	0.13	0.19	0.25	0.18	0.24	0.37	0.28	0.21	0.20	0.30	0.40	0.25	0.28	0.35
36	0.05	0.30	0.15	0.14	0.31	0.20	0.16	0.19	0.25	0.25	0.23	0.13	0.12	0.21	0.21	0.20	0.28	0.33	0.11	0.22	0.30	0.39	0.27	0.24	0.43
37	0.05	0.12	0.03	0.07	0.20	0.04	0.02	0.03	0.07	0.08	0.03	0.04	0.06	0.04	0.06	0.16	0.17	0.04	0.07	0.14	0.22	0.14	0.11	0.21	
38	0.16	0.14	0.08	0.20	0.35	0.31	0.18	0.12	0.27	0.21	0.19	0.17	0.21	0.27	0.19	0.28	0.48	0.38	0.20	0.22	0.41	0.46	0.36	0.48	0.43
39	0.20	0.24	0.15	0.26	0.34	0.36	0.30	0.18	0.32	0.30	0.25	0.22	0.26	0.23	0.29	0.41	0.38	0.28	0.17	0.24	0.37	0.43	0.36	0.31	0.37
40	0.11	0.15	0.00	0.15	0.33	0.31	0.15	0.08	0.21	0.15	0.12	0.14	0.21	0.23	0.14	0.27	0.35	0.29	0.16	0.20	0.39	0.42	0.24	0.25	0.34
41	0.21	0.19	0.26	0.32	0.35	0.35	0.35	0.29	0.34	0.46	0.46	0.40	0.42	0.46	0.47	0.45	0.42	0.40	0.46	0.51	0.35	0.55	0.30	0.53	0.38
42	0.23	0.20	0.31	0.32	0.37	0.48	0.26	0.26	0.38	0.42	0.47	0.40	0.40	0.46	0.45	0.43	0.44	0.34	0.48	0.35	0.51	0.30	0.47	0.31	
43	0.25	0.21	0.15	0.26	0.46	0.41	0.23	0.18	0.34	0.31	0.33	0.21	0.36	0.39	0.30	0.38	0.50	0.49	0.33	0.58	0.41	0.54	0.40	0.36	
44	0.20	0.30	0.27	0.31	0.35	0.44	0.26	0.21	0.33	0.40	0.39	0.33	0.44	0.47	0.37	0.52	0.47	0.42	0.38	0.38	0.43	0.68	0.48	0.40	0.41
45	0.15	0.22	0.22	0.20	0.26	0.40	0.26	0.19	0.34	0.34	0.30	0.30	0.29	0.31	0.41	0.41	0.39	0.32	0.39	0.35	0.59	0.49	0.37	0.41	0.43
46	0.17	0.24	0.16	0.26	0.43	0.42	0.21	0.19	0.34	0.31	0.28	0.32	0.32	0.36	0.45	0.41	0.52	0.51	0.30	0.35	0.49	0.57	0.44	0.45	0.45

SIMILARITIES BETWEEN MUTUAL FUND PORTFOLIOS

[illegible]

Table A5

ALTERNATIVE M-D-SCAL SOLUTIONS--MARCH 31, 1971

Mutual Fund	One Dimensional Solution Coordinate	Two Dimensional Solution Coordinates		Three Dimensional Solution Coordinates		
		X	Y	X	Y	Z
1	-2.029	-1.723	-0.691	-1.327	-0.800	-0.741
2	-1.566	-0.003	-1.512	-0.154	-0.182	-1.441
3	-2.231	-1.470	-1.215	-0.952	-1.170	-0.906
4	-1.272	-0.941	-0.866	-0.738	-0.307	-0.956
5	-0.023	-0.228	0.480	-0.346	0.590	-0.164
6	-0.281	0.183	-0.643	0.261	-0.602	-0.387
7	-1.388	-0.523	-1.359	-0.593	-1.258	0.441
8	-1.733	-1.206	-1.032	-1.187	-0.676	-0.718
9	-0.382	-0.759	0.333	-0.163	0.331	-0.857
10	-0.616	-0.644	-0.307	-0.696	-0.303	-0.096
11	-0.953	-0.724	-0.690	-0.616	-0.763	-0.375
12	-1.077	-1.082	-0.443	-1.081	-0.190	-0.347
13	-0.807	-0.516	-0.693	-0.419	-0.620	-0.518
14	-0.451	-0.523	-0.026	-0.587	-0.128	0.171
15	-0.848	-0.924	-0.245	-0.843	-0.551	0.196
16	-0.316	-0.349	-0.137	-0.374	-0.283	0.137
17	0.242	0.173	0.311	0.128	0.422	-0.111
18	0.045	0.012	0.038	-0.073	0.160	-0.064
19	-0.758	-1.152	0.215	-0.828	-0.446	0.811
20	-0.641	-0.572	-0.345	-0.571	-0.516	0.041
21	0.256	0.027	0.479	-0.100	0.500	0.254
22	0.190	0.192	0.057	-0.175	-0.028	0.161
23	0.360	0.056	0.711	-0.083	0.651	0.433
24	-0.129	-0.291	0.064	-0.344	0.002	0.164
25	0.747	0.602	0.650	0.448	0.204	0.829
26	-0.089	-0.305	0.193	-0.304	-0.109	0.496
27	0.397	-0.193	1.107	-0.374	0.904	0.619
28	-0.003	-0.427	0.639	-0.515	0.269	0.701
29	0.435	0.396	0.118	0.358	0.059	0.288
30	0.552	0.140	0.895	0.000	0.873	0.449
31	0.736	0.731	0.223	0.704	0.208	0.203
32	0.788	0.173	1.085	-0.024	1.085	-0.001
33	0.538	0.472	0.332	0.421	0.355	0.182
34	0.818	0.759	0.197	0.716	0.308	0.002
35	1.556	1.390	0.201	1.322	0.174	0.151
36	1.264	1.182	0.003	1.133	0.094	-0.105
37	2.971	2.364	0.236	2.133	0.223	0.121
38	1.180	1.058	0.361	0.987	0.377	0.192
39	1.026	0.984	0.007	0.956	0.100	-0.019
40	1.557	1.377	0.211	1.311	0.207	0.126
41	0.144	0.228	-0.130	0.212	-0.291	0.244
42	-0.223	0.045	-0.471	0.103	-0.583	-0.160
43	0.531	0.592	0.062	0.603	0.069	-0.012
44	0.208	0.165	0.067	0.138	0.016	0.088
45	0.664	0.662	-0.044	0.622	-0.175	0.331
46	0.611	0.593	0.193	0.561	0.199	0.145

HIERARCHY OF NATURAL CLUSTERS IN THE TWO DIMENSIONAL M-D-SCAL MUTUAL FUND SOLUTION SPACE

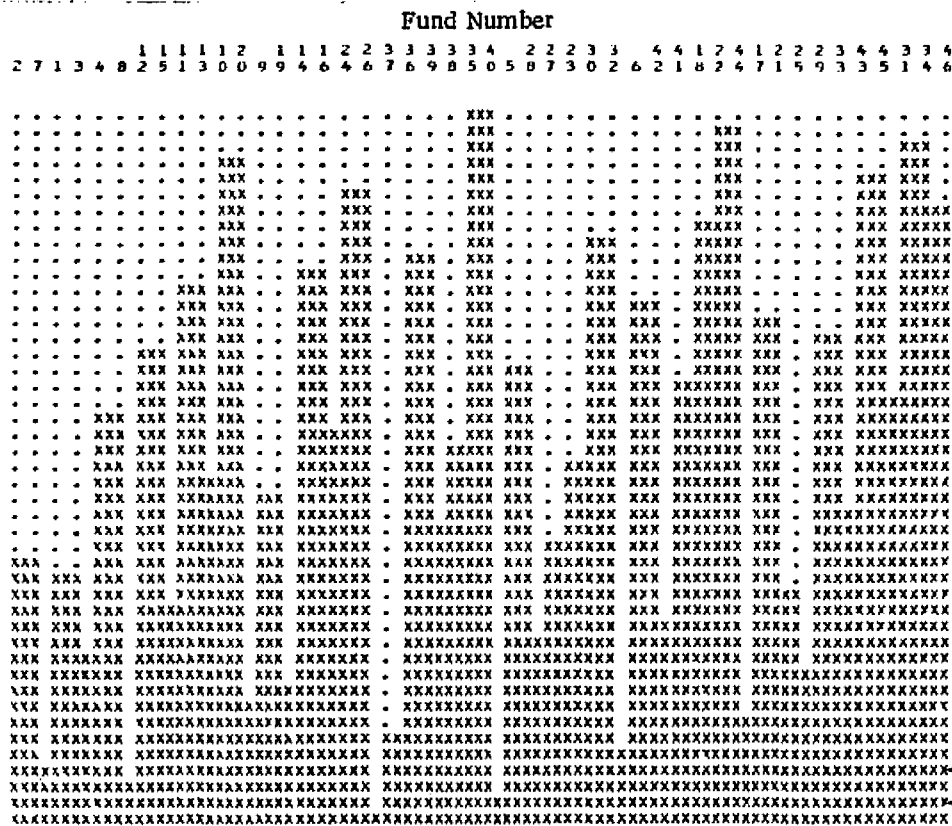


Figure A6 (b)

HIERARCHY OF NATURAL CLUSTERS IN THE THREE DIMENSIONAL M-D-SCAL MUTUAL FUND SOLUTION SPACE

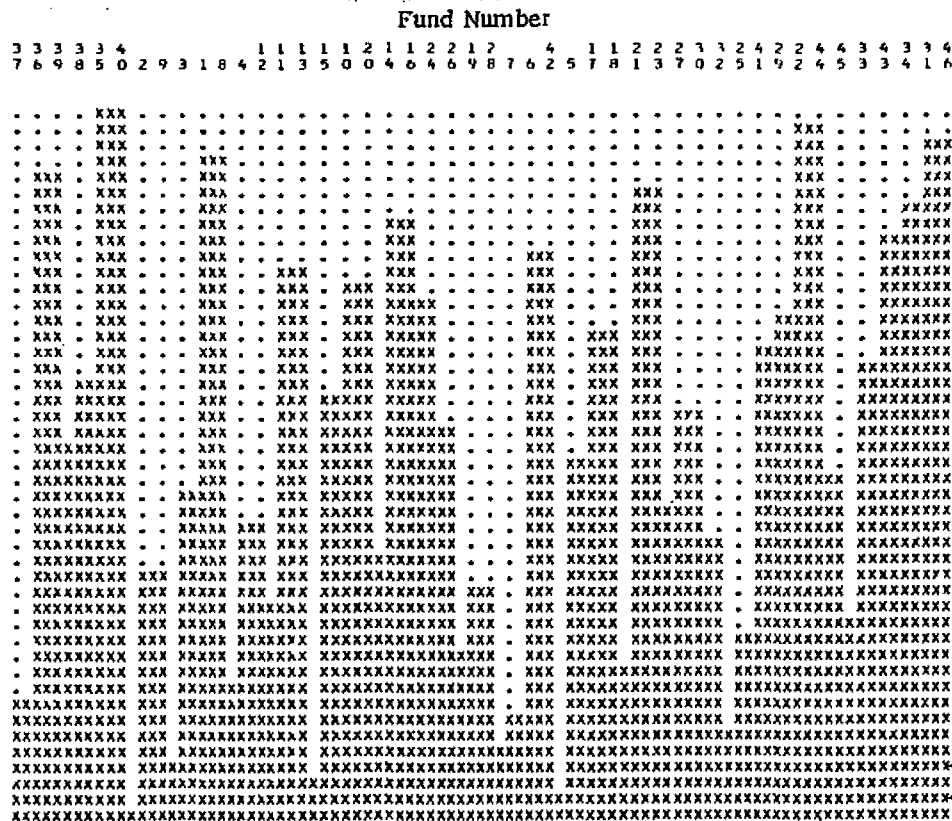


Table A7

MUTUAL FUND YIELD AND GROWTH PROPERTY VALUES

Mutual Fund	Percent Annual Yield	Percent Growth In Share Value				
		1/1/68 to 3/31/71	3/31/71 to 6/30/71	3/31/71 to 9/30/71	3/31/71 to 12/31/71	3/31/71 to 3/31/72
1	2.50	- 9.90	0.42	-6.52	-2.06	14.23
2	1.40	-23.23	-2.90	-5.00	-0.95	13.47
3	0.50	-39.37	6.38	4.25	11.70	27.72
4	1.60	-17.25	-2.25	-2.92	1.15	6.13
5	3.00	7.26	1.54	2.18	6.58	16.28
6	2.10	-31.38	0.65	-1.39	-0.01	10.08
7	0.70	-38.99	-1.73	-2.42	-5.01	3.65
8	0.60	-13.04	2.36	2.55	9.46	31.35
9	0.60	4.39	-0.61	1.52	3.04	10.63
10	0.60	-16.34	0.68	0.85	5.22	16.14
11	1.40	21.32	-0.86	-0.29	5.90	16.91
12	1.50	3.43	4.17	5.27	10.47	23.23
13	1.60	-30.07	6.53	8.54	18.46	30.58
14	2.40	- 5.95	0.65	-1.98	2.48	13.99
15	1.60	17.81	5.03	5.52	11.68	22.64
16	2.30	2.38	-2.04	-2.27	0.46	8.16
17	2.70	22.69	-1.25	-4.09	0.44	9.81
18	3.10	17.27	-0.94	-0.60	4.24	8.43
19	2.10	12.62	2.61	3.30	7.73	23.89
20	2.00	5.23	4.34	4.96	8.51	18.18
21	3.50	2.72	-0.09	0.65	1.29	4.01
22	3.20	19.26	-1.74	-0.33	3.11	8.88
23	3.80	12.61	-0.08	-2.72	-0.59	4.33
24	2.90	14.56	-0.82	-1.92	-0.02	5.48
25	4.90	4.90	-2.48	-3.89	-1.28	2.98
26	2.70	28.76	3.36	5.06	8.57	13.44
27	3.40	26.93	0.62	0.92	2.69	5.64
28	3.20	21.19	0.84	0.84	4.28	11.24
29	3.80	6.05	-1.52	-1.26	2.04	6.38
30	3.90	27.73	-0.59	-0.12	0.68	3.06
31	4.60	21.14	-1.49	0.68	4.73	7.30
32	3.10	24.57	-0.12	-1.09	2.60	7.83
33	4.40	17.03	-1.23	-0.46	2.23	7.16
34	4.50	22.77	-1.81	-1.25	3.57	5.56
35	5.50	7.64	-0.90	1.91	5.40	10.07
36	5.40	15.55	-3.34	-4.27	-8.09	-3.57
37	6.80	19.13	-0.90	2.26	7.07	9.95
38	4.70	33.15	-1.42	0.89	3.66	8.84
39	5.20	31.56	-1.13	1.59	6.95	13.64
40	4.90	18.93	-1.24	1.27	3.21	7.08
41	3.20	21.04	0.05	2.03	5.74	10.28
42	2.70	14.82	1.26	2.64	8.08	18.26
43	4.10	11.30	0.08	2.20	5.69	9.20
44	3.40	14.10	-2.73	-3.41	0.97	4.05
45	3.20	14.65	0.70	2.29	6.17	11.92
46	4.20	28.58	-1.86	-0.72	-1.17	1.15

Table A8

ALTERNATIVE M-D-SCAL SOLUTIONS--SEPTEMBER 30, 1971

Mutual Fund	Two Dimensional Solution Coordinates		Three Dimensional Solution Coordinates		
	X	Y	X	Y	Z
1	-1.366	-0.891	-0.859	-0.316	-1.176
2	0.091	-1.376	-0.040	-0.284	-1.304
3	-1.161	-1.506	-0.869	-1.205	-0.841
4	-1.165	-0.165	-0.918	-0.026	-0.638
5	-0.541	0.282	-0.512	0.433	-0.266
6	-0.193	-0.649	-0.115	-0.499	-0.518
7	-1.326	0.269	-1.202	0.469	-0.350
8	-1.222	-1.078	-1.118	-0.829	-0.624
9	0.185	-0.653	0.038	-0.733	0.593
10	-0.774	-0.404	-0.783	-0.313	-0.335
11	-0.767	-0.715	-0.720	-0.861	0.103
12	-1.072	-0.242	-1.079	-0.165	-0.145
13	-0.773	-1.007	-0.615	-0.952	-0.500
14	-0.478	-0.124	-0.511	-0.271	0.271
15	-0.934	-0.383	-0.842	-0.580	0.490
16	-0.211	-0.185	-0.248	-0.305	0.242
17	0.113	0.173	0.106	0.202	-0.155
18	-0.074	0.059	-0.138	0.153	-0.145
19	-0.858	0.698	-0.546	-0.264	1.032
20	-0.513	-0.298	-0.554	-0.424	0.141
21	0.012	0.411	-0.055	0.480	0.082
22	0.136	-0.007	0.141	-0.100	0.131
23	0.224	0.673	0.126	0.692	0.403
24	-0.377	0.113	-0.432	0.008	0.404
25	0.786	0.994	0.561	0.632	-0.897
26	-0.260	0.366	-0.250	0.014	0.708
27	-0.152	0.880	-0.249	0.903	0.193
28	-0.389	0.455	-0.460	0.342	0.463
29	0.206	0.289	0.163	0.309	0.110
30	0.007	0.815	-0.105	0.821	0.255
31	0.801	0.161	0.775	0.203	0.032
32	0.040	1.376	-0.072	1.302	0.456
33	0.647	0.288	0.620	0.314	0.118
34	0.973	0.263	0.952	0.249	0.175
35	1.178	-0.007	1.160	-0.018	-0.030
36	1.123	0.261	1.046	0.284	-0.126
37	2.359	0.289	2.059	0.254	0.109
38	1.088	0.266	1.040	0.299	0.008
39	1.088	0.179	1.054	0.237	0.026
40	1.347	0.304	1.293	0.262	0.135
41	0.233	-0.059	0.203	-0.146	0.448
42	-0.091	-0.583	-0.049	-0.744	0.055
43	0.620	0.041	0.624	-0.008	0.043
44	0.346	0.139	0.319	0.072	0.241
45	0.542	0.046	0.540	-0.084	0.382
46	0.551	0.241	0.540	0.193	0.204

Table A9

ALTERNATIVE M-D-SCAL SOLUTIONS--JUNE 30, 1971

Mutual Fund	Two Dimensional Solution Coordinates		Three Dimensional Solution Coordinates		
	X	Y	X	Y	Z
1	-1.027	-0.457	-0.492	-0.231	-0.951
2	-0.407	-1.632	-0.229	-0.278	-0.520
3	-1.540	-1.152	-0.976	-1.203	-0.822
4	-1.304	-0.128	-1.060	-0.090	-0.655
5	-0.425	0.250	-0.480	0.329	-0.053
6	-0.063	-0.619	-0.008	-0.575	-0.460
7	-1.174	0.302	-0.955	0.400	-0.587
8	-1.486	-0.837	-1.149	-0.759	-0.783
9	0.164	-0.978	-0.134	-0.889	0.652
10	-0.706	-0.128	-0.637	-0.211	-0.310
11	-0.626	-0.874	-0.584	-0.922	-0.042
12	-1.180	0.026	-1.177	-0.068	-0.217
13	-0.987	-0.808	-0.711	-0.896	-0.484
14	-0.415	-0.133	-0.468	-0.271	0.164
15	-0.743	-0.586	-0.838	-0.614	0.345
16	-0.223	-0.193	-0.282	-0.264	0.244
17	0.248	0.266	0.252	0.270	-0.136
18	0.028	0.202	-0.031	0.295	-0.045
19	-0.177	-0.699	-0.493	-0.311	0.974
20	-0.461	-0.339	-0.559	-0.390	0.081
21	0.027	0.457	-0.047	0.538	0.124
22	0.145	-0.016	0.109	-0.068	0.109
23	0.092	0.719	0.005	0.678	0.471
24	-0.394	0.033	-0.499	-0.003	0.375
25	0.410	0.855	0.437	0.371	-0.768
26	-0.202	-0.005	-0.289	-0.092	0.510
27	-0.140	0.878	-0.168	0.918	0.181
28	-0.390	0.574	-0.403	0.288	0.773
29	0.317	0.345	0.259	0.378	0.140
30	-0.012	0.842	-0.095	0.843	0.379
31	0.726	0.288	0.695	0.302	0.134
32	-0.557	1.209	-0.661	1.135	0.259
33	0.739	0.376	0.701	0.375	0.213
34	0.682	0.277	0.660	0.259	0.181
35	1.213	0.206	1.183	0.100	0.089
36	1.178	0.382	1.121	0.349	-0.037
37	2.484	0.452	2.177	0.301	0.177
38	1.103	0.366	1.072	0.299	0.083
39	1.061	0.198	1.017	0.172	-0.101
40	1.409	0.363	1.321	0.305	0.160
41	0.267	-0.075	0.219	-0.159	0.320
42	0.070	-0.556	0.025	-0.724	0.089
43	0.679	0.152	0.675	0.085	0.132
44	0.275	0.050	0.221	0.025	0.200
45	0.725	-0.086	0.723	-0.213	0.192
46	0.592	0.234	0.552	0.215	0.217

Three Dimensional Solution Coordinate

Mutual Fund	X		Y		X	Y	Z
	X	Y	X	Y			
1	-0.701	-1.295	-0.622	-0.545	-1.117		
2	0.313	-1.259	0.006	-0.071	-1.269		
3	-1.176	-0.994	-0.867	-1.055	-0.483		
4	-0.953	-0.565	-0.877	-0.265	-0.561		
5	-0.477	0.332	-0.539	0.383	-0.150		
6	0.113	-0.696	0.119	-0.709	-0.360		
7	-1.494	-0.383	-1.022	0.119	-0.968		
8	-1.356	-1.060	-1.233	-0.952	-0.373		
9	-0.090	-0.772	-0.124	-0.548	0.949		
10	-0.813	-0.578	-0.784	-0.428	-0.434		
11	-0.759	-0.451	-0.652	-0.746	0.227		
12	-1.257	-0.253	-1.204	-0.175	-0.224		
13	-0.755	-0.769	-0.590	-0.857	-0.321		
14	-0.602	-0.194	-0.577	-0.390	0.249		
15	-0.929	-0.039	-0.784	-0.477	0.530		
16	-0.275	-0.072	-0.264	-0.245	0.434		
17	0.135	0.277	0.087	0.347	-0.177		
18	-0.061	0.047	-0.137	0.142	-0.166		
19	-1.030	0.575	-0.669	-0.157	1.047		
20	-0.483	-0.185	-0.486	-0.362	0.302		
21	0.184	0.666	0.094	0.730	0.222		
22	0.093	0.037	0.074	-0.031	0.155		
23	0.233	0.785	0.115	0.855	0.169		
24	-0.395	0.252	-0.433	0.153	0.356		
25	1.172	0.352	0.993	0.415	-0.412		
26	-0.259	0.431	-0.227	0.148	0.698		
27	-0.072	0.791	-0.176	0.839	-0.067		
28	-0.575	0.611	-0.642	0.507	0.359		
29	0.266	0.232	0.226	0.282	0.076		
30	0.011	0.876	-0.111	0.906	0.160		
31	0.840	0.181	0.802	0.202	-0.048		
32	-0.740	1.525	-0.818	1.367	0.242		
33	0.663	0.263	0.633	0.283	0.031		
34	0.947	0.289	0.938	0.219	0.104		
35	1.149	-0.115	1.101	-0.163	-0.166		
36	1.243	0.211	1.176	0.239	-0.118		
37	2.242	0.220	1.992	0.082	-0.024		
38	1.115	0.169	1.062	0.155	-0.078		
39	0.950	0.122	0.925	0.111	-0.101		
40	1.363	0.260	1.274	0.206	0.052		
41	0.147	0.013	0.153	-0.100	0.273		
42	0.024	-0.495	0.075	-0.717	0.147		
43	0.650	0.068	0.647	0.018	-0.048		
44	0.332	0.275	0.289	0.177	0.386		
45	0.556	0.060	0.564	-0.110	0.332		
46	0.510	0.258	0.493	0.220	0.167		

Table A11

PROPERTY FIT OF INVESTMENT GROUP VECTORS

Investment Group	Number of Funds Invested in Group (46 Maximum)	Correlation Between Fitted and Original Vector	Direction Cosines of Fitted Vectors with Respect to Dimension	
			X	Y
1	22	0.2685	-0.5089	-0.8608
2	18	0.3651	0.5119	0.8590
3	29	0.3256	-0.6700	-0.7424
4	15	0.4739	0.3792	0.9253
5	30	0.2453	0.1180	0.9930
6	38	0.6591	-0.2637	-0.9646
7	20	0.3791	-0.9957	-0.0928
8	19	0.1558	0.9182	-0.3961
9	9	0.5676	-0.9588	-0.2840
10	32	0.2118	0.6442	-0.7649
11	14	0.5026	-0.9594	-0.2819
12	14	0.1065	0.9849	0.1732
13	16	0.3442	-0.8122	0.5833
14	15	0.5406	0.0952	0.9955
15	30	0.3878	0.4509	0.8626
16	12	0.4960	-0.2132	0.9770
17	7	0.3715	-0.9005	0.4348
18	7	0.5684	0.3008	0.9537
19	23	0.5349	-0.9380	0.3465
20	18	0.2932	-0.4334	-0.9012
21	32	0.3124	-0.8352	-0.5499
22	18	0.1989	-0.5424	-0.8401
23	23	0.5007	0.1428	0.9898
24	8	0.6288	-0.5588	0.8293
25	18	0.5396	-0.5949	-0.8038
26	23	0.0944	0.1688	-0.9857
27	12	0.7386	0.0683	-0.9977
28	13	0.5428	0.5072	-0.8618
29	9	0.5825	0.4623	0.8867
30	6	0.4328	-0.4352	-0.9003
31	9	0.0420	-0.0101	-0.9999
32	15	0.2979	-0.9495	0.3136
33	11	0.6263	-0.9798	0.2000

(Continued)

Table A11 (Continued)

Investment Group	Number of Funds Invested in Group (46 Maximum)	Correlation Between Fitted and Original Vector	Direction Cosines of Fitted Vectors with Respect to Dimension	
			X	Y
34	13	0.4175	0.4673	0.8841
35	15	0.6818	-0.9973	-0.0730
36	35	0.3657	-0.6952	-0.7189
37	10	0.8667	0.5699	0.8217
38	17	0.3934	-0.9543	0.2988
39	10	0.5042	0.1033	-0.9946
40	18	0.4466	-0.2949	-0.9555
41	24	0.0667	-0.7011	0.7131
42	14	0.6506	0.0207	-0.9998
43	4	0.9987	-0.6559	-0.7549
44	29	0.3501	-0.9467	0.3220
45	8	0.3800	-0.6796	-0.7336
46	24	0.1784	-0.9954	0.0959
47	9	0.5799	-0.3806	0.9247
48	4	0.9999	0.8090	0.5878
49	5	0.1802	0.4342	0.9008
50	3	0.1000	0.8236	0.5672
51	15	0.7109	0.3921	0.9199
52	29	0.4120	0.1069	-0.9943
53	23	0.4405	-0.4602	-0.8878
54	21	0.5964	-0.2854	0.9584
55	9	0.5629	0.6961	0.7179
56	29	0.1923	-0.2617	-0.9652
57	41	0.5322	-0.0317	-0.9995
58	9	0.8914	0.4603	0.8878
59	22	0.0427	-0.3984	-0.9172
60	17	0.6634	-0.0296	-0.9996
61	17	0.2820	-0.4444	0.8958
62	19	0.5320	-0.2067	0.9784
63	9	0.2769	0.9816	0.1912
64	6	0.0757	0.5645	0.8254
65	23	0.3698	0.7003	0.7139

(Continued)

Table A11 (Continued)

Investment Group	Number of Funds Invested in Group (46 Maximum)	Correlation Between Fitted and Original Vector	Direction Cosines of Fitted Vectors with Respect to Dimension	
			X	Y
66	9	0.3652	-0.9023	-0.4311
67	37	0.2745	-0.8217	-0.5700
68	11	0.2896	0.2043	0.9789
69	11	0.5227	0.4772	0.8788
70	21	0.4605	0.1773	-0.9842
71	17	0.2074	-0.4081	-0.9129
72	7	0.9113	0.1006	0.9949
73	18	0.2802	-0.1939	-0.9810
74	5	0.4776	-0.2013	0.9795
75	4	0.8636	-0.9079	-0.4193
76	6	0.8465	-0.9139	0.4059
77	6	0.6418	-0.9742	-0.2259
78	15	0.4262	-0.2106	0.9776
79	9	0.7882	0.8807	0.4737
80	24	0.1745	0.1855	0.9826
81	14	0.2856	-0.0518	-0.9987
82	8	0.6846	-0.9912	0.1322
83	5	0.7651	0.2793	0.9602
84	21	0.5850	-0.9984	-0.0562
85	27	0.5964	-0.1902	-0.9817
86	8	0.9030	0.6838	0.7297
87	12	0.3985	0.1178	-0.9930
88	5	0.6269	0.2979	0.9546
89	11	0.4715	-0.8477	-0.5304
90	36	0.4692	-0.0108	-0.9999
91	33	0.0385	0.5395	-0.8420
92	27	0.5379	-0.2846	-0.9586
93	5	0.5328	-0.6467	-0.7627
94	13	0.5473	0.2604	0.9655
95	46	0.2136	0.4476	-0.8942
96	38	0.9025	0.9599	0.2803
97	39	0.5407	0.6923	-0.7217

Table A12(a)

MDPREF SOLUTION--STIMULI POINTS
INVESTMENT GROUP COORDINATES

Investment Group	Coordinates		Investment Group	Coordinates	
	X	Y		X	Y
1	0.0016	-0.0059	25	-0.0356	-0.0051
2	-0.0229	-0.0353	26	0.0066	-0.0100
3	0.0593	0.0825	27	-0.0322	-0.0584
4	-0.0412	-0.0369	28	-0.0376	-0.0404
5	0.1012	0.2399	29	-0.0485	-0.0400
6	0.1654	0.1938	30	-0.0499	-0.0671
7	-0.0063	0.0875	31	-0.0463	-0.0487
8	-0.0260	-0.0251	32	-0.0337	-0.0098
9	-0.0455	-0.0093	33	-0.0446	-0.0170
10	0.0391	0.0012	34	-0.0455	-0.0303
11	-0.0351	-0.0170	35	-0.0213	0.0246
12	-0.0297	-0.0309	36	0.0988	0.2327
13	-0.0305	-0.0041	37	-0.0485	-0.0574
14	-0.0331	-0.0044	38	-0.0270	-0.0299
15	0.0106	0.0163	39	-0.0393	-0.0427
16	-0.0236	0.0490	40	-0.0171	0.0186
17	-0.0455	-0.0219	41	-0.0020	-0.0003
18	-0.0513	-0.0335	42	-0.0398	-0.0489
19	-0.0030	0.1073	43	-0.0512	-0.0578
20	-0.0231	-0.0034	44	0.1783	0.4827
21	0.0379	0.1028	45	-0.0491	-0.0298
22	-0.0173	0.0230	46	-0.0142	0.0160
23	-0.0068	0.0289	47	-0.0501	-0.0509
24	-0.0530	-0.0583	48	-0.0529	-0.0586

(Continued)

Table A12(a) (Continued)

Investment Group	Coordinates		Investment Group	Coordinates	
	X	Y		X	Y
49	-0.0492	-0.0566	74	-0.0488	-0.0655
50	-0.0548	-0.0589	75	-0.0565	-0.0597
51	-0.0258	0.0376	76	-0.0526	-0.0440
52	0.0432	0.0294	77	-0.0553	-0.0492
53	0.0048	0.0148	78	-0.0336	-0.0173
54	0.0015	0.0762	79	-0.0457	-0.0563
55	-0.0454	-0.0395	80	-0.0208	-0.0228
56	0.0205	0.0500	81	0.0061	-0.0161
57	0.2554	0.2506	82	-0.0503	-0.0276
58	-0.0487	-0.0522	83	-0.0528	-0.0494
59	0.0052	-0.0305	84	-0.0038	0.0798
60	-0.0205	-0.0317	85	0.0298	0.0080
61	-0.0185	-0.0365	86	-0.0459	-0.0469
62	-0.0297	-0.0402	87	-0.0289	-0.0506
63	-0.0414	-0.0453	88	-0.0549	-0.0549
64	-0.0512	-0.0465	89	-0.0403	-0.0302
65	0.0165	0.0063	90	0.1309	0.0714
66	-0.0482	-0.0372	91	0.0409	0.0369
67	0.0507	0.1187	92	0.0023	-0.0303
68	-0.0458	-0.0358	93	-0.0510	-0.0332
69	-0.0321	0.0322	94	-0.0391	-0.0147
70	-0.0079	-0.0009	95	0.3084	0.2678
71	-0.0185	0.0084	96	0.7640	-0.5514
72	-0.0484	-0.0209	97	0.1911	-0.0329
73	-0.0244	-0.0197			

Table A12(b)

MDPREF SOLUTION--SUBJECT VECTORS
MUTUAL FUND DIRECTION COSINES

Mutual Fund	Direction Cosines With Respect to Dimension		Mutual Fund	Direction Cosines With Respect to Dimension	
	X	Y		X	Y
1	0.8311	0.5561	24	0.6448	0.7644
2	0.8589	0.5122	25	0.9984	-0.0574
3	0.2628	0.9648	26	0.7373	0.6756
4	0.4378	0.8991	27	0.7609	0.6489
5	0.7569	0.6536	28	0.6897	0.7241
6	0.9907	-0.1358	29	0.9903	-0.1388
7	0.7030	0.7112	30	0.9387	0.3449
8	0.1296	0.9916	31	0.9147	-0.4042
9	0.7236	0.6902	32	0.8800	0.4750
10	0.4097	0.9122	33	0.9681	-0.2504
11	0.5627	0.8267	34	0.9621	-0.2725
12	0.3188	0.9478	35	0.8508	-0.5255
13	0.6579	0.7531	36	0.8828	-0.4697
14	0.5139	0.8578	37	0.8231	-0.5680
15	0.4222	0.9065	38	0.8879	-0.4600
16	0.6053	0.7960	39	0.8830	-0.4694
17	0.9967	-0.0812	40	0.8760	-0.4823
18	0.9441	0.3295	41	0.9922	-0.1243
19	0.4401	0.8980	42	0.9819	0.1894
20	0.6285	0.7778	43	0.9392	-0.3435
21	0.9142	0.4053	44	0.9917	0.1286
22	0.9984	-0.0558	45	0.9660	-0.2586
23	0.9354	0.3536	46	0.9405	-0.3398

BIBLIOGRAPHY

- Beckenbach, E. and Bellman, R. An Introduction to Inequalities. New York: Random House, 1961.
- Broad Street Investing Corporation, prospectus. New York: May 1, 1971.
- Carroll, J. Douglas. "Individual Differences and Multidimensional Scaling." in Multidimensional Scaling: Theory and Applications in the Behavioral Sciences; Volume I, Theory, ed. R. N. Shepard, et. al.; New York: Seminar Press, 1972.
- Chang, J. J. and Carroll, J. D. "How to Use MDPREF, a Computer Program for Multidimensional Analysis of Preference Data." mimeographed report (Bell Telephone Laboratories, Murray Hill, N. J., October, 1969).
- Chang, J. J. and Carroll, J. D. "How to Use PROFIT, a Computer Program for Multidimensional Analysis of Preference Data." mimeographed report (Bell Telephone Laboratories, Murray Hill, N. J., October, 1969).
- Clarkson, Geoffrey P. E. Portfolio Selection: A Simulation of Trust Investment. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1962.
- Cohen, Jerome B. and Zinbarg, Edward D. Investment Analysis and Portfolio Management. Homewood, Ill.: Richard D. Irwin, Inc., 1967.
- Cole, L. C. "The Measurement of Partial Interspecific Association." Ecology, XXXVIII (1957), 226-233.
- Coombs, C. H. A Theory of Data. New York: John Wiley & Sons, Inc., 1964.
- Deese, James. "On the Structure of Associative Meaning." Psychological Review, LXIX (May, 1962), 161-175.
- Farrar, Donald E. The Investment Decision Under Uncertainty. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1962.
- Feeney, A. and Hester, D. "Stock Market Indices: A Principal Components Analysis." Risk Aversion and Portfolio Choice. Edited by D. Hester and J. Tobin. New York: John Wiley & Sons, Inc., 1967.

- Fisher, L. and Lorie, J. "Some Studies of Variability of Returns on Investments in Common Stocks." Journal of Business, XLIII (April, 1970).
- Friend, Irwin; Brown, F. E.; Herman, Edward S.; and Vickers, Douglas. "A Study of Mutual Funds," prepared for the S. E. C. Report of the House Committee on Interstate Commerce & Finance, 87th Congress, 2nd Session. U. S. Government Printing Office, 1962.
- Friend, Irwin; Blume, Marshall; and Crockett, Jean. "Mutual Funds and Other Institutional Investors: A New Perspective." A Twentieth Century Fund Study, 1970.
- Friend, Irwin and Vickers, Douglas. "Portfolio Selection and Investment Performance." Journal of Finance, XX (September, 1965), 391-415.
- Green, Paul E. and Carmone, Frank J. Multidimensional Scaling and Related Techniques in Marketing Analysis. Marketing Science Institute Series. Boston: Allyn and Bacon, Inc., 1970.
- Green, Paul E. and Maheshwari, Arun. "Common Stock Perception and Preference: An Application of Multidimensional Scaling." Journal of Business, XLII (October, 1969), 439-457.
- Green, Paul E. and Rao, Vithala. Applied Multidimensional Scaling: A Comparison of Approaches and Algorithms. New York: Holt, Rinehart and Winston, Inc., 1972.
- Green, Paul and Rao, Vithala. "A Note on Proximity Measures and Cluster Analysis." Journal of Marketing Research, (August, 1969), 359-364.
- Gregson, R. A. M. "Quadratic Similarities." British Journal of Mathematical and Statistical Psychology, XXIII (May, 1970), 53-68.
- Harshman, Richard A. "PARAFAC." Working Papers in Phonetics No. 16. University of California, Los Angeles, December, 1970.
- Harshman, Richard A. "PARAFAC 2." Working Papers in Phonetics No. 22. University of California, Los Angeles, March, 1972.
- Ilano, Alberto R. "Alternative Approaches to Risk in Portfolio Analysis." unpublished Ph. D. dissertation. Berkeley, Calif.: University of California, 1969.

- Jensen, Michael D. "The Performance of Mutual Funds in the Period 1945-1964." Journal of Finance, XXIII (May, 1968), 389-416.
- Johnson, Stephen C. "Hierarchical Clustering Schemes." Psychometrika, XXXII (September, 1967), 241-254.
- King, Benjamin. "Market and Industry Factors in Stock Price Behavior." Journal of Business, XXXIX (January, 1966), 139-190.
- Kruskal, Joseph and Carmone, Frank. "How to Use M-D-SCAL, A Program to Do Multidimensional Scaling and Multidimensional Unfolding (Version 5M of MDSCAL all in Fortran IV)." mimeographed report (Bell Telephone Laboratories, Murray Hill, N. J., October, 1969).
- Arthur Lipper Corporation. "INSIGHT." four books. New York, 1971.
- Arthur Lipper Corporation. "Mutual Fund Performance Analysis." a weekly bulletin. New York, March 31, 1971.
- Markowitz, Harry M. "Portfolio Selection." Journal of Finance, VII (March, 1952), 77-91.
- Markowitz, Harry M. Portfolio Selection: Efficient Diversification of Investments. New York: John Wiley & Sons, Inc., 1959.
- Meyers, Stephen L. "A Factor Analysis Approach to Studying the Structure of the Stock Market." unpublished Ph. D. dissertation. Philadelphia: University of Pennsylvania, 1970.
- "Mutual Fund Group Classifications." FundScope, April, 1971, 200-201.
- Pinches, George E. and Kinney, William R., Jr. "The Measurement of Volatility of Common Stock Prices." Journal of Finance, XXVI (March, 1971), 119-125.
- Restle, Frank. Psychology of Judgment and Choice. New York: John Wiley & Sons, Inc., 1961.
- Rummel, Rudolph J. Applied Factor Analysis. Evanston, Ill.: Northwestern University Press, 1970.
- Sharpe, William F. "A Linear Programming Algorithm for Mutual Fund Portfolio Selection." Management Science, XIII (March, 1967), 499-510.
- Sharpe, William F. "Mutual Fund Performance." Journal of Business, XXXIX (January, 1966), 119-138.

- Shepard, Roger N. "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function - Part I." Psychometrika, XXVII (June, 1962), 125-140; Part II, XXVII (September, 1962), 219-246.
- Shepard, Roger N. "A Taxonomy Of Some Multidimensional Scaling Methods," mimeographed paper prepared for the Bell-Penn Workshop on Multidimensional Scaling. Philadelphia: University of Pennsylvania, June 7-10, 1972.
- Smith, Keith V. Portfolio Management. New York: Holt, Rinehart and Winston, Inc., 1971.
- Smith, Keith V. and Schreiner, John C. "A Portfolio Analysis of Conglomerate Diversification." Journal of Finance, XXIV (June, 1969), 413-427.
- Sokal, R. and Sneath, P. Principle of Numerical Taxonomy. San Francisco: W. H. Freeman and Company, 1963.
- Standard & Poor's Corporation. "Analysts Handbook." New York, published annually.
- Standard & Poor's Corporation. "Mutual Fund Summary." Stock Guide, (March, 1971), 235-249.
- Torgerson, W. S. Theory and Methods of Scaling. New York: John Wiley & Sons, Inc., 1960.
- True, D. and Matson, R. "Cluster Analysis and Multidimensional Scaling of Archeological Sites in Northern Chile." Science, CLXIX, No. 3951 (1970), 1201.
- Tucker Ledyard R. "Relationship of Scaling and Factor Analysis." mimeographed paper delivered before the Bell-Penn Workshop on Multidimensional Scaling. Philadelphia: University of Pennsylvania, June 7-10, 1972.
- United States Bureau of the Budget, "Standard Industrial Classification Manual." U. S. Government Printing Office, 1971.
- United States Securities and Exchange Commission, "Directory of Companies Filing Annual Reports with the Securities and Exchange Commission." U. S. Government Printing Office, June, 1969.
- United States Securities and Exchange Commission. "Institutional Investor Study Report of the Securities and Exchange Commission." report to the Committee on Interstate and Foreign Commerce, 92nd Congress, 1st Session. 8 volumes, U. S. Government Printing Office, 1971.

Von Neuman, John and Morgenstern, Oskar. Theory of Games and Economic Behavior. 3rd ed. Princeton, N. J.: Princeton University Press, 1953.

Wagenaar, W. and Padmos, P. "Quantitative Interpretation of Stress in Kruskal's Multidimensional Scaling Technique," British Journal of Mathematical and Statistical Psychology, XXIV (1971) 101-110.

Whitehall Fund, Inc., prospectus. New York. May 1, 1971.

Wiesenberger Service, Inc. Investment Companies 1971. New York, 1971.

Wish, Myron and Carroll, J. Douglas. "Applications of 'INDSCAL' to Studies of Human Perception and Judgment," mimeographed report (Bell Telephone Laboratories, Murray Hill, N. J., 1971), to be published in Handbook of Perception, ed. M. P. Friedman, New York: Academic Press, 1973.

Wish, Myron. "Notes on the Variety, Appropriateness, and Choice of Proximity Measures." mimeographed paper prepared for the Bell-Penn Workshop on Multidimensional Scaling. Philadelphia: University of Pennsylvania, June 7-10, 1972.

Young, Forrest W. "Bibliography of Multidimensional Scaling Programs Developed at the University of Michigan, Bell Telephone Laboratories and the University of North Carolina Including Condensed Descriptions and Abstracts of Unpublished Programs." mimeographed paper prepared for the Bell-Penn Workshop on Multidimensional Scaling. Philadelphia: University of Pennsylvania, June 7-10, 1972.

Young, Forrest W. "TORSCA, an IBM Program for Nonmetric Multidimensional Scaling." Journal of Marketing Research, V (August, 1968), 319-321.