Winter 1994

An inquiry into non-survey techniques for updating input-output coefficients: Comparative experiments with data from the Soviet Union

Ali Reza Jalili
University of New Hampshire, Durham

Order Number 9518482

An inquiry into non-survey techniques for updating input-output coefficients: Comparative experiments with data from the Soviet Union

Jalili, Ali Reza, Ph.D.

University of New Hampshire, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

AN INQUIRY INTO NON-SURVEY TECHNIQUES FOR UPDATING INPUT-OUTPUT COEFFICIENTS: COMPARATIVE EXPERIMENTS WITH DATA FROM THE SOVIET UNION

BY

ALI REZA JALILI

BS, N.I.O.C COLLEGE OF ACCOUNTANCY AND FINANCE, 1976

MBA, JAMES MADISON UNIVERSITY, 1979

MA, UNIVERSITY OF NEW HAMPSHIRE, 1983

DISSERTATION

SUBMITTED TO THE UNIVERSITY OF NEW HAMPSHIRE

IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

ECONOMICS

DECEMBER, 1994

This dissertation has been examined and approved.

Marc W. Herold

Dissertation Chair, Marc W. Herold
Associate Professor of Economic Development

Richard W. England

Richard W. England
Professor of Economics

Evangelos Simos

Evangelos O. Simos
Professor of Economics

Dennis Meadows

Dennis Meadows
Professor of Systems Management

William H. Crown

William Crown
Senior Researcher, Brandeis University

August 26, 1994
Date

TO ALL THOSE WHO RELENTLESSLY STRUGGLE

FOR

FREEDOM, EQUALITY, AND HUMAN DIGNITY

TO MY PARENTS,

MAJDALDIN JALILI AND OZRA ETEMADI,

WHO TAUGHT ME THE MEANING OF SACRIFICE

TO HEVA AND KIKI,

WHO EACH IN HER OWN WAY,

ENDURED IMMENSELY FOR THIS

ACKNOWLEDGEMENTS

No person is solely the result of individual effort. The outcome is always the culmination of efforts by numerous individuals. This is particularly true in my case. I cannot articulate my sentiments more aptly than Einstein, who once said "many times a day I realize how much my own outer and inner life is built upon the labors of my fellow-men, both living and dead, and how earnestly I must exert myself in order to give in return as much as I have received." Obviously, I can never give back as much as I received. For I have been fortunate beyond expectation in my journey. I have been blessed by having so many dear friends who are my sanctuaries. My path came to cross those of so many wise, able, and giving individuals who granted me so much, so graciously, and so selflessly. I am forever indebted, and pay homage to each and every one of them. Limitation of space, however, compels me to mention only a small subset of the group as a token of my tribute to the entire roster.

First and foremost, I want to enunciate my deepest heartfelt gratitude to my friend and mentor Dr. Ahmad Afrasiabi, one of those rare individuals who combines depth and breadth of knowledge with a deep love for humanity, and his lovely family, Kristin and Cyrus. They opened their hearts to me at the time I needed it most, and, like a true family, unconditionally supported and showered me with all kinds of tenderness. Thank you Behrooz, Kris, and Cyrus.

This would have not been possible without you. Special thanks also goes to my advisor, mentor, and chair of the committee, Dr. Marc Herold, who was a tremendous help every step of the way and from whom I learned enormously during the years. Further, I would like to offer my sincere appreciation to members of my dissertation committee. Dr. William Crown of Brandeis University, who showed me grace; Dr. Richard England, Dr. Evangelos Simos, and Dr. Dennis Meadows, who were tolerant, conducive, and understanding far beyond expectation. I am also honored to mention Dr. Sam Rosen, who showed me that economists can be compassionate. Dr. Richard Mills, who trained me to use mathematics in economics; The late Dr. Joseph Kosnik, who was the manifestation of gentleness; Dr. Lester Bittle; who taught me how to synthesize. Dr. Farhad Noomani and Dr. Iraj Hoshi, who introduced me to economics; and Mr. Nejat Khosh-Ahang, who made me to be at peace with mathematics. A note of recognition is also due to my colleagues at Allegheny College and Mercer University, whose collegiality and support enabled me to conclude this task.

Lastly, my thanks to my siblings Zohreh, Marzieh, Muhammad, and Hamid, who always stood by me; and my beloved friends; Massoud, Farkhondeh, Babak, Hamed, Naser, Kambiz, Hosein, Farhad, Amir, Rahman, Habib, Farah, Farzad, and Keywan; An oasis of hope in the desert of turbulence; friends who always believed in and never gave up on me.

TABLE OF CONTENTS

CHAPTER PAGE

CHAPTER PAGE

LIST OF TABLES

LIST OF TABLES CONTINUED

LIST OF TABLES CONTINUED

LIST OF TABLES CONTINUED

LIST OF TABLES CONTINUED

LIST OF TABLES CONTINUED

LIST OF TABLES CONTINUED

LIST OF TABLES CONCLUDED

LIST OF FIGURES

LIST OF FIGURES CONTINUED

LIST OF FIGURES CONTINUED

LIST OF FIGURES CONTINUED

FIGURE PAGE

LIST OF FIGURES CONCLUDED

ABSTRACT

AN INQUIRY INTO NON-SURVEY TECHNIQUES FOR UPDATING INPUT-OUTPUT COEFFICIENTS: COMPARATIVE EXPERIMENTS WITH DATA FROM THE SOVIET UNION

by

ALI REZA JALILI

UNIVERSITY OF NEW HAMPSHIRE, DECEMBER, 1994

The study begins with explanation of I-O tables, intertemporal stability of its coefficients, and logic of updating techniques. Following a literature review, nine non-survey updating methods are selected and utilized to update the actual 1966 table of the Soviet Union to the target year of 1972. Next, the data and simulation procedure are specified and justified. The concept of matrix comparison along with methods to accomplish this task are discussed. Then, the resultant updated matrices are compared with the actual data, via employment of 25 criterions. Accordingly, RAS and Friedlander procedures are ranked as top performers. The results, while reasonable in holistic sense, are not impressive partitively.

Exogenous estimation of a subset of coefficients is considered next. Several possibilities for "selective targeting" are investigated, and three such criterions, namely "key," "most important," and "largest" coefficients are adopted. These criterions, then are used to modify RAS,

Friedlander, and NAIVE methods via incorporation of exogenous data. An additional approach, Residual Minimum method, is also employed. Thus, ten "modified" estimates are obtained and compared with the actual table. The outcome indicates substantial improvements in the RAS and Friedlander updates, particularly when exogenously estimated subset consists of the largest, or the "most important" coefficients. NAIVE method displays reasonable improvement. Inclusion of prior information, however, in some instances leads to deterioration of individual estimates for the remaining coefficients.

Finally, sectoral aggregation and its effects on the performances of updating methods are addressed. Aggregated estimates at three levels of sectoral details, i.e., 36, 16, and 6 sectors, are obtained and comparisons are performed. The results indicate that, generally the performances of the updating methods, as well as the intertemporal stability of coefficients, are direct functions of the level of sectoral detail. No change in ranking occurs due to aggregation.

Conclusions of this research may be used for selection of updating methods, as well as in construction phase of tables, to identify and focus on the most influential coefficients. Throughout, a rather detailed presentation of methods and statistical tools are offered. All experiments are conducted for both direct and inverse matrices. NAIVE, (constant coefficient), is added for comparative purposes.

CHAPTER ONE

INTRODUCTION

> "From the arch of the bridge to which his guide has carried him, Dante now sees the Diviners.....coming slowly along the bottom of the forth chasm. By help of their incantations and evil agents, they had endeavored to pry into the future which belongs to He almighty alone, and now their faces are painfully twisted the contrary way; and being unable to look before them, they are forced to walk backwards."
>
> Dante Alighirie
> *Divine Comedy: The Inferno*

1.1) BACKGROUND:

In any given economy there exist an interrelationship, an interdependence, amongst various industries. Input-Output (I-O) analysis refers to a quantitative framework for studying this relationship. I-O analysis, in its present form has been developed by Wassily Leontief (Leontief 1936, 1941, and other works) who was awarded the Nobel prize in Economic Science in 1973 in recognition of his path breaking works in this field.

The historical origin of inter-industry relationships, however, can be traced back to 1758 and Francios Quesney's work "*Tableau Economique*." There, Quesney depicted the relationship among four social classes of his time and the circulation of wealth amongst them, i.e., the general interdependence of different sectors in the economy. In 1847, Leon Walras published his work "*Elements d'economie politique pure*," which was another attempt in quantitative representation of the sectoral interdependence within an economy. He presented this interrelationship in the context of a set of simultaneous linear equations linking producing and consuming sectors. The model was capable of demonstrating the effects of any change in one equation on the entire system. Walras was concerned with simultaneous determination of all prices in an economy. In doing so, he shifted the economists' focus from neo-classical's "partial equilibrium" back to classical's "general equilibrium." At the same time, and in addition to the general equilibrium of exchange, he paid attention to the general equilibrium of production and spoke about "coefficients of production." His system not only capture the interdependence among the producers and their respective competing demands for various input factors, but also includes equations for consumer income and expenditure. The demand for and supply of various products and inputs, sectoral costs of production, and the consumers' ability to substitute different commodities are also included in his model. Thus, his system can legitimately be considered as a major step in development of I-O analysis.

Karl Marx's "simple" and "expanded" reproduction schemes, developed in "*Das Kapital*," can be viewed as another step in the same direction. In his model, the output of "department one," that is capital goods, becomes the inputs for "department two," viz., consumer goods. The product of the latter reenters the production process as the inputs for the former. Hence the two sectors in the economy are linked through this model and their interrelationship can be studied. In fact this model can be regarded as highly aggregated (two sectors) Input-Output model.

Contributions to the theory of general equilibrium by the Swedish economist Gustav Cassel and the Italian economist Vilferdo Pareto were also among significant steps in the evolution of I-O analysis. In resent history, the works of two Soviet economists, Popov and Groman can be cited. Using 1923-24 data for the Soviet Union, they constructed a table showing the interconnection among different sectors of the economy. They sought to determine that for production of any given quantity of any given commodity, how much of the other commodities are needed.

They called their model "Balance of The National Economy. This model, to use Alec Nove's phrase, can be called the grandfather of the modern Input-Output tables. Their work, however, was not pursued in the Soviet Union. After Bukharin and his "Balanced Growth" idea were discredited, the works of Popov and Groman was also dismissed, since it was concerned with "balance growth." Other early Soviet economists working on I-O analysis include Chayanov, Bogdanov, and Kristman (see Belykh 1989).

It remained for Wassily Leontief and a later date to finish the task. His works (1936, 1941), formulated and presented the I-O analysis in its modern format. Leontief, who started contemplating the analysis of interindustry relationships from the early years of his career, found not much enthusiasm and support for his ideas in the Soviet Union. Later, while still a young economist, he went to Western Europe and then migrated to the United States. In the U.S he continued his work and provided the culmination of a long journey by furnishing a comprehensive quantitative model of interindustry relationships. His ideas launched a new sub-field in the economic discipline and triggered a series of studies, elaborations, applications, extension, and refinements that continues to date.

The I-O analysis today encompasses a wide range of topics and has numerous applications. It provides policy makers and economic planners with an extremely powerful tool for policy simulation, impact analysis, as well as host of other uses. A host of national, regional, sectoral, and corporate analysis can be conducted utilizing I-O analysis, resulting in deeper understanding of potential problems and their possible sources. Therefore problems may be avoided before they even arise, or at least a better preparation in facing them may be facilitated by the virtue of being aware of their existence.

1.2) PURPOSE AND PLAN OF THE PRESENT STUDY:

The construction of I-O tables, by their very nature, demands an elaborate and accurate statistical apparatus as well as a well trained personnel. Even under best of conditions, a rather substantial time lag exists between the time of actual census and construction and publication of the survey based table. Time lags of five to ten years are not uncommon. Construction costs through actual surveys are rather high and in some instances prohibitive. The situation is exacerbated for regional as well as less developed countries due to budget constraints and inadequacy of appropriate personnel and organization. Moreover, I-O analysis assumes constancy of technical coefficients over time. Many researchers in applied field, however, are concerned over this assumption and raise the question of the validity of the results obtained while relying on the constancy assumption. It is generally accepted that the coefficients do change over time and the longer the time lag the larger these changes are expected to be. This in turn, makes the aforementioned concerns even more acute. If survey based tables were readily available, these concerns would have been minimal. But construction of I-O survey based tables will encounter the problems mentioned above. For these reasons and to facilitate the utilization of I-O analysis, researchers have attempted to device some "short cuts" in order to construct a table without actually going through a survey. These attempts have led to the emergence of a variety of "non-survey" techniques of constructing an I-O table. Each one of these techniques, however, has its own shortcomings and difficulties, which raise the question of their usefulness and applicability. In other words, given the shortcomings and operational requirements of these techniques, would the results obtained through them be any more accurate than accepting the assumption of constancy of coefficients? Would the loss of accuracy be justified by cost and time gains? In short would the remedy be worse than the disease? The present study intends to address some of these and related issues.

Following the introduction, a brief theoretical presentation of I-O analysis along with the underlying assumptions and their associated concerns will be presented in chapter two. Chapter three, will take up the issue of temporal stability of the I-O coefficients as well as the non-survey techniques of constructing an I-O table by presenting a survey of the literature on the subject and probing the major scholarly works. Chapter four will describe nine different estimating and updating techniques selected in this study for the purpose of empirical work. The chapter examines theoretical foundations and mathematical presentations of these methods, along with their shortcomings and criticism. Chapter five explains implementation of the project, i.e., the methodology, simulation procedures, and the data. The chapter also addresses data accuracy and compatibility issues and explains the steps taken to transform the data into a workable format for the proposed research. Providing the concepts of matrix comparisons, as well as related shortcomings and difficulties associated with these comparisons, along with statistical tests and criterions used for such comparisons, concludes chapter five. Presentation of the results of the experiments along with their interpretation and explanation is the subject of chapter six.

A concern among I-O analysts is necessity of inclusion of additional exogenous information into the updating process, viz., modification of minimum information techniques, along with selection criterion and treatment of such information. Chapter seven, will cover these issues and provides four different modified techniques and three schemes for selection of exogenous information, leading to ten different modified estimates of the 1972 Soviet table. The resultant tables and their evaluations and comparisons are covered in chapter eight.

The I-O researchers in many instances find themselves forced to work with aggregated tables. The effect of aggregation and possible bias introduced as the result is a great concern for many scholars in the field. Involvement in this aspect of I-O analysis is outside the scope of the current study. However, as a matter of interest in analyzing the effects of aggregation on the performance of various updating techniques, chapter nine attempts, without getting involved in the theoretical issues, to explore this area by applying selected updating techniques to the Soviet tables at three different levels of sectoral aggregation and report the results. Chapter ten, summarizes major findings and states the limitations of the present study. Concluding remarks, along with possible avenues and suggestions for further research completes chapter ten as well as the project.

CHAPTER TWO

THEORETICAL FOUNDATION

"The best way to be boring is to leave nothing out."

Voltaire

Even though input-output analysis is well known, a brief discussion of the method, its foundations, and its assumptions seems warranted, not only to expedite comprehension of issues raised here, but to establish a common nomenclature as well. Thus, in the following pages, input-output analysis along with relevant theoretical basis and underlying assumptions will be briefly explained. To reaffirm the cognizance, a hypothetical table also will be given, the table's mathematical relationships will be explained and presented, and the connection of I-O tables to the national accounts will be demonstrated.

2.1) EXPLANATION OF THE TABLE:

In the words of William Miernyk (1965) the I-O analysis broadly speaking is part of economics statistics, or the field of econometrics. It is a theory of production based on the notion of interdependence among various economic agents and institutions. It can also be viewed as simplification of Walrasian general equilibrium, for it aggregates Walras' "commodities" into "outputs," thereby significantly reducing their numbers. The analysis relies exclusively on the Input-Output table. The table is constructed from actual economic data for a given geographical unit in a given time period. Conceptually, the table can have as many "sectors" as there are commodities. Practically, however, the choice is limited by availability of data and resources. A "sector" can range from a single firm to a group of rather diverse enterprises combined together. It can be constructed in monetary, physical, or other units of measurement. Although physical unit is a more accurate reflection of each sector's usage of the output of other industries, there will be measurement problems when the aggregations are such that more than one commodity fall under a given industry, e.g., a mainframe high speed computer and a basic small personal computer may both be classified in one industry and thus be counted as two units, despite their enormous price differential. Adoption of monetary value is not free of problems either, and can introduce some measurement difficulties due to the possibility of price changes without any actual variation in real physical usage of any given input. As a matter of convention, ease of data collection, computation, and comparison, monetary tables are utilized more frequently than any other variant.

The table itself consists of four quadrants. The upper left hand quadrant is variously termed "interindustry transaction table," "the processing sector," or "the producers sector." This quadrant contains all the producers of goods and services in the economy and shows the flow of products from each sector (as a supplier) to other sectors (as consumers). The rows of this table, then, depict the distribution of each sector's output throughout the economy (i.e., the breakdown of each sector's output sales) and the columns show the composition of inputs required by each industry to produce its output (i.e., the breakdown of each sector's input purchases). The lower left-hand quadrant is called "the payments sector" or "the value added sector." Included in this quadrant are nonindustrial payments in the economy for a given time period, e.g., wages, interest, profit, payments to the government, gross inventory depletion, imports, and depreciation allowances. The third quadrant or the upper right segment of the table is referred to as "the final demand sector" and shows the breakdown of the sales of each industry to its final customers. This is the "autonomous" section of the I-O table. Any change in this sector will be transmitted to other parts of the table. Included in this sector are personal consumption expenditures, gross private domestic investment, government purchases, exports, and gross inventory accumulation. The forth and final quadrant is referred to as "total gross output" or "total gross outlays," or- with some modification in arrangement of the data- "gross national product." The row sum of the table will yield "total gross outlays" and the column sum will show the "total gross output" for each sector. These sums, of course, must be equal for each individual sector as well as the table as a whole. A typical table is given below for demonstrative purposes. This table contains (n) producing sectors (which make up the interindustry relationships or the first quadrant). The total interindustry demand and supply of inputs (the column sum and row sums of the inter-industry transaction table) are given by U and V. There are four final demand sectors in the table, namely consumption, gross investment, government purchases, and exports. The third quadrant includes the import, as well as the value added shares of government and household sectors. The row and column sum totals are denoted by (X) and represent total gross outlays and total gross outputs respectively. These sums, are obviously equal.

The connection between national income and national product is apparent from the table. Row sums of the last

TABLE 2-1

A HYPOTHETICAL INPUT-OUTPUT TABLE

TO FR	1	2	.	.	n	U	C	I	G	E	X
1	x_{11}	x_{12}	.	.	x_{1n}	u_1	c_1	i_1	g_1	e_1	X_1
2	x_{21}	x_{22}	.	.	x_{2n}	u_2	c_2	i_2	g_2	e_2	X_2
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
n	x_{n1}	x_{n2}	.	.	x_{nn}	u_n	c_n	i_n	g_n	e_n	X_n
V	v_1	v_2	.	.	v_n	-	-	-	-	-	-
M	m_1	m_2	.	.	m_n	-	m_c	m_i	m_g	-	M
T	t_1	t_2	.	.	t_n	-	t_c	t_i	t_g	t_e	T
H	h_1	h_2	.	.	h_n	-	h_c	h_i	h_g	h_e	H
X	x_1	x_2	.	.	x_n	-	C	I	G	E	H

column will yield:

(2-1) $$X = X_1 + X_2 +.+.+ X_n + M + T + H$$

And the column sums of the last row will give:

(2-2) $$X = X_1 + X_2 +.+.+ X_n + C + I + G + E$$

Setting the two equations equal and simplifying, results in:

(2-3) $$C + I + G + (E\text{-}M) = T + H$$

Where C, I, G, E, and M represent consumption, investment, government expenditure, export, and import respectively. Thus, (2-3) indicates equality of National Income and National Product.

2.2) ASSUMPTIONS UNDERLYING INPUT-OUTPUT ANALYSIS:

Input-output theory is fundamentally a theory of production. As such, it contains the following implicit and explicit assumptions about production:

1) Each industry produces only one commodity as its output and each output is classified only under one industry.
2) The economy can be divided into various sectors, and each sector can encompass many producers.

3) Inputs are not substitutable.
4) At any point in time only one technology is utilized for production of any given commodity.
5) The production function exhibits constant return to scale, which implies absence of economies or diseconomies of scale.
6) sectors use their required inputs in fixed proportion.

Of course the plausibility of each of these assumptions can be debated. The first two assumptions, however, do not present a theoretical obstacle to input-output theory. A systematic and elaborate classification scheme as well as a more disaggregated table can satisfy these assumptions. The third and forth assumptions imply that their imposition is either technologically warranted or the constancy of the relative prices renders any alteration of the production process and inputs useless. So, if substitution among inputs is possible under a given sectoral classification, it is conceptually possible to generate a table so detailed that guarantees inapplicability of input substitution for any given sector at any given time. Furthermore it should be noted that inputs in the production process are very often highly complementary. This fact, at least in the short run, greatly reduces input substitution even in the case of a change in relative prices. With regards to the possibility of existence of more than one production technique, it can be argued that even in such cases, the profit maximization behavior of firms will lead to adoption of only one of these possibilities.

Given these, the other assumptions will be acceptable as well, viz., a linear homogeneous production function with constant return to scale as well as constant set of technical coefficients.

Since the intertemporal stability of coefficients is the most relevant of the I-O assumptions to the current research, it merits further discussion. First, one should investigate the factors that conceptually could lead to changes in the coefficients. Next step, then, is to design methods of measuring the changes and empirically verify the effects of these factors on the I-O coefficients. If such changes could be verified, then one is faced with the problem of capturing these changes and incorporating them into the table in an appropriate manner to transform it into a usable and accurate tool of economic policy.

There can be many factors affecting the coefficients. Some of the most important ones are mentioned below:

-Change in relative prices can lead to input substitution and thereby altere technological coefficients.

-The international disparity among input prices or any other such factor can cause some industries to switch from foreign suppliers to domestic ones or visa-versa, hence changing the coefficients.

-Technological change which can cause change of inputs or their required quantity.

-The change in product mix of each sector.

-Presence of economies or diseconomies of scale.

Theoretically, any one or any combination of the above factors can change the I-O coefficients. As the time lag between the base year and forecasting year widens, the probability of such change increase. The question, however, is how fast and to what extend? As was mentioned earlier, even in the presence of these factors, there are some plausible explanations for stability of the coefficients, at least in the short and medium runs. Thus the issue becomes one of empirical verification, which will be dealt with in the following chapters.

2.3) MATHEMATICAL PRESENTATION:

Dividing the economy into (n) sectors and denoting the total output and the total final demand of the i-th sector as $X_{ij}$and Y_{ij} respectively, the following relationship can be written:

(2-4) $$X_i = z_{i1} + z_{i2} + \ldots\ldots + z_{i.n} + Y_i$$

or:

(2-5) $$X_i = \sum_{j=1}^{n} z_{ij} + Y_i$$

where z_{ij} denotes the flow of i-th sector's output to other sectors, viz., the interindustry sales by sector (i). The right hand side, then, is the sum of all total final demands for the i-th sector's output plus its interindustry sales.

The equation (2-5) can be generalized to include all sectors, that is:

(2-6) $$\sum_{i}^{n} X_i = \sum_{i}^{n}\sum_{j}^{n} z_{ij} + \sum_{i}^{n} Y_i$$

Given the assumptions of I-O analysis, one can write:

(2-7) $$a_{ij} = \frac{z_{ij}}{X_j}$$

Which basically indicates that the flow of input from (i) to (j) entirely and exclusively depends on the j-th sector's level of production.

In the equation(2-7) X_j is the total output of the sector (j), and a_{ij} is the constant "technical coefficient" (also known as "input-output coefficient" or "direct input coefficient"). Utilizing (2-7), the equation (2-6) can be rewritten as:

(2-8) $$\sum_{i}^{n} X_i = \sum_{i}^{n}\sum_{j}^{n} a_{ij} X_j + \sum_{i}^{n} Y_i$$

for i=1, 2, ..., n and j=1, 2,...., n

In this equation Y_i's and a_{ij}'s are known quantities and X_i's are to be determined. Denoting the vector of X_i's (total gross outputs) and the vector of Y_i's (total final demands) as [X] and [Y] respectively and the (n) by (n) matrix of technical coefficients (a_{ij}'s) as [A] and bringing all the X terms to the left-hand side, one can write the system of equations represented by (2-8) in matrix notation as:

(2-9) $$(I - A)(X) = Y$$

where [I] is an (n) by (n) identity matrix. This, of course, makes [I-A] an (n) by (n) matrix with $(1-a_{11})$, $(1-a_{22})$,...., ..., $(1-a_{nn})$ on its main diagonal and $-a_{ij}$ everywhere else. Solving for X from (2-9) will yield:

(2-10) $$X = (I - A)^{-1} (Y)$$

The existence of a unique solution in this system depends on singularity of [I-A]. That is, existence or nonexistence of the "Leontief Inverse," $[I - A]^{-1}$, which in turn depends on whether the determinant of the matrix [I-A] is equal to zero or not. In other words the equation system represented by (2-10) will have a unique solution if the determinant of $[I - A] \neq 0$, hence guaranteeing the existence of the Leontief inverse and non-singularity of [I-A]. The relation of gross output of each sector to each of the final demands is apparent from (2-10). Denoting the elements of Leontief inverse by α_{ij}, this dependence may be shown as

$$\frac{\partial X_i}{\partial Y_j} = \alpha_{ij}$$

In the equation system (2-10), [Y] cannot be negative since a negative value here implies a negative demand, which is an economic absurdity. The values of [X] also cannot be negative, for it implies another logical contradiction, i.e., negative production. In other words not only the conditions for a unique solution must be met, but also the requirement that the unique solution must not be negative should be satisfied. So, if [Y] is to generate non-negative values for [X] in (2-10), all elements of the Leontief matrix must be non-negative. Since $0 \leq a_{ij} < 1$ by definition, then if the determinant of $[I - A] > 0$, non-negativity of all elements of the Leontief matrix is assured. In other words the Hawking-Simon conditions for this system imply strictly positive principal minors for $[I-A]$.

Following this brief introduction to input-output tables, the assumption of constancy (or stability) over time of the I-O coefficients should be explored further. For if these coefficients are stable, one might not need to resort to updating techniques to construct usable interim tables.

If, on the other hand, coefficients are intertemporally unstable, the need for non-survey estimation methods becomes apparent. Next chapter explores the literature in pursuit of these questions and prepares the stage for subsequent empirical verifications.

CHAPTER THREE

REVIEW OF THE LITERATURE

"When you take stuff from one writer, it's plagiarism; but when you take it from many writers, it's research."

Wilson Mizner

3.1) INTRODUCTION:

A central question in input-output analysis, is the question of intertemporal stability of technical coefficients. It is concern over this issue that has absorbed an enormous amount of the researchers' time and energy. Many studies have been conducted to investigate the validity of constancy assumption, discovering the potential sources of change in the coefficients, exploring the ramification of these changes for I-O analysis, devising techniques and methods to overcome the problems arising from these changes, etc. The question is particularly crucial for researchers engaged in applied work. Because if the coefficients are not constant, the accuracy and reliability of the results, hence the implied recommendations, will be seriously questioned.

Therefore, first and foremost a researcher must determine whether the I-O coefficients are constant over time, or at least reasonably stable, hence reliable for research or not. Theoretical issues aside, the emanating question is how fast and to what extent the technical coefficients do change over time, if at all, and what are the consequences of these (possible) changes, i.e., the problem is actually an empirical one.

3.2) INTERTEMPORAL STABILITY OF I-O COEFFICIENTS:

Numerous studies, utilizing a variety of methods, have attempted to address these concerns, and there are several ways to verify the constancy of the technical coefficients. For instance if two (or more) actual tables are available, suppose for time t_0 and t_1, then by determining the product of the Leontief matrix of one (say t_0) and the inverse of the Leontief matrix of the other ($t_1$in this case) the constancy of the coefficients can be verified. For if, (3-1) $$[I-A_{t_0}]\,[I-A_{t_1}]^{-1}=I$$

then:

(3-2) $$A_{t_o}=A_{t_1}$$

hence indicating no structural changes between the two periods. Similarly the same result can be obtained if the product of the transaction matrix of one period and the inverse of the transaction matrix of another period is obtained. If this product is equal to [I], constancy of the coefficients may be concluded, if the equality does not hold the difference between the product and the identity matrix can reveal cell by cell changes that have taken place in the structure of the economy during the respective time period.

Another way to examine the constancy assumption used by researchers such as Carter (1970) and Miller and Blair (1985) is to plot a set of coefficients at two time periods, t_0 and t_1 on a graph with same scale on both axis. Then, if there were no changes in the coefficients, all of the plotted observations will fall on a forty five degree line through the origin. Any deviation from this line represents a change in the coefficients.

Carter(1970), utilizing these techniques, examined the structural changes of the U.S. economy during 1947-1958 period. She found that for some industries the coefficients do not demonstrate a great deal of change, while others do so. Those which changed in some cases exhibited decline while some showed increase. The changes were partially explained by input substitution during the period.

Another approach used by Carter(1970) is to use the actual coefficients of a given year, say t_0 , along with the final demands of a later period, say t_1, and calculate the gross output that would have been required to meet t_1's final demands if there were no changes in the technical coefficients. Then by comparing these estimations with the actual outputs the constancy can be verified. Her study used this method for the final demand of the year 1961 along with coefficients of the years 1939, 1947, and 1958, and found that for the most part the I-O coefficients are remarkably stable over time.

A study by Tilanus and Rey (1964) for the Netherlands' tables during 1948-1958 basically supported these findings. Leontief himself(1953), using the 1939 coefficients and 1919 as well as 1929 final demands of the U.S economy, generated the gross output requirements of those years. He then compared these figures with the results obtained through two other approaches and concluded the superiority of the I-O generated estimates. Since I-O predictions were made with the assumption of constancy of the coefficients and were superior to the results obtained via other methods, this was viewed as an indication of validity of the constancy assumption. However, a similar study by Barnet (1954) which utilized the 1939 U.S input-output table, employed regression analysis along with other methods and concluded that regression results are more accurate than those obtained via I-O analysis, hence indicating changes in the coefficients.

The constancy tests are also conducted for specific industry or regions. For example, three survey based tables of the state of Washington for 1963, 1967, and 1972 have been used in order to determine the stability of the regional coefficients. Beyers (1972) conducted one such study. On individual, cell by cell or sector by sector, basis the results were mixed. On the whole, however, the deviations were not large enough to impair the analysis. For instance the 1967 total output requirements and intermediate output requirement derived by utilization of 1963 coefficients along with 1967 final demand, overestimated the actual 1967 total output and intermediate output requirements by 2.3% and 10.5% respectively. The deviation on some individual sectors, however, were as high as 77%.

Another study by Conway (1980) arrived at similar conclusions. Using 1963, 1967, and 1972 data for the state of Washington, he found that total output of 1972 on the basis of 1963 and 1967 coefficients applied to 1972 final demand vector was overestimated 0.8% in case of 1963 and underestimated by 1.2% in case of 1967. Here again, while overall results indicate stability of the coefficients, some individual cells or industries performed poorly.

Using 1965 and 1970 data for the state of Kansas, Emerson (1976) arrived at the same mixed conclusion. Working at the sectoral level, Baster (1980) found that during 1974-1976 period over 90% of the trade coefficients for the Strathclyde region of Scotland were stable. The Same study revealed that at a firm level, 79% of the coefficients were constant and 13.5% had variations under 10%.

As suggested by Miller and Blair (1985), "backward linkage" and "forward linkage" can also be used as summary measures of comparison of coefficients over time. The direct backward linkage, that is $\sum_{i=1}^{n} a_{ij}$ or the sum of the elements of the j-th column of the transaction matrix, shows the dependence of this sector for inputs on various other sources. Forward linkage, that is $\sum_{j=1}^{n} a_{ij}$ or the sum of the elements of the i-th row of the transaction matrix, reveals dependence of other sectors of the economy on the output of sector (i).

Now, in comparing two time periods (0) and (t) if $\sum_{i=1}^{n} a_{ij}^{0} < \sum_{i=1}^{n} a_{ij}^{t}$ it can be concluded that the sector (j) has grown more dependent on other sectors in the economy during the period under study. If in the comparison, one finds $\sum_{j=1}^{n} a_{ij}^{0} < \sum_{j=1}^{n} a_{ij}^{t}$,then it can be concluded that other sectors have grown more dependent on the i-th sector as supplier of their input. Of course "total" backward or forward linkages (i.e., output or input multipliers) can be easily determined by calculating the sum of the elements in j-th column or i-th row of the Leontief inverse.

This approach has been used by Conway (1977) for the state of Washington based on actual tables of 1963, 1967, and 1972. He also found rather stable overall coefficients (average multipliers were 1.33, 1.29. and 1.33 for the years 1963, 1967, and 1972 respectively) while some individual sectors showed deviation of up to 18 percent.

Another study by Conway (1975) for the state of Washington using 1963 and 1967 final demands ($i'Y_{63}$ and $i'Y_{67}$) and 1963 output $(i'X_{63})$ estimated the 1967 output ($i'X_{67}$) using an approach known as "final demand blowup. Formally this can be written as:

$$i'X_{67} = [i'X_{63}]\left[\frac{i'Y_{67}}{i'Y_{63}}\right] \qquad (3\text{-}3)$$

The result was an overestimation of the 1967 output by 3.1 percent. This result was inferior to the one obtained through I-O analysis, but was obtained through a much simpler method and at the same time was not that far off from the I-O results (2.3% overestimation). However, as noted by Miller and Blair (1985), these close results may exist at a highly aggregated level, but they do not have the I-O analysis' ability to produce detail results at the sectoral level.

Other notable studies in the area of input-output coefficient changes include Evans (1954), Theil (1957), vaccara (1969), Sevaldson (1970), Simonovits (1975), Bezdek (1984), and Sawyer (1992). Review of this literature reveals that even though the I-O coefficients are relatively stable, they do gradually change over time. They might be appropriate for short and medium term "impact analysis," but long run forecasts based on them can not be reliable. The claim of inappropriateness of constant coefficient assumption for long term analysis basically relies on the fact that technology and relative prices do change over long run. The literature also reveals that, the stability is more apparent at the aggregate level, and the errors at this level are not large. But in the mid to long range analysis and at the sectoral level, the I-O analysis do produce rather large errors.

The reason for more accuracy at higher aggregation level and overall table is relatively straightforward. As one increases the level of aggregation in an I-O table, one includes more and more products in one sector. Then, changes in some sectors may cancel one another and thereby reducing the overall deviation. The deviations in opposite direction are due to switching and substitution. Moreover, when more products are combined in one sector, the relative weight of each product in that sector decreases, hence a given change in any sector will not have the same consequences on the aggregated coefficients as it may have on the coefficients of a more disagregated one.

3.3) NON-SURVEY UPDATING TECHNIQUES:

In recognition of intertemporal instability of I-O coefficients, researchers directed their attention to devising techniques that are capable of capturing these changes and incorporating them in the I-O tables. This derive led to construction of methods of updating and projecting I-O coefficients, i.e., the body of literature that is generally referred to as "partial survey" and "non-survey" techniques. In the final analysis, all these methods attempt to turn the fixed coefficient I-O model, which is a one point observation of the economy, into a model suitable and reliable for forecasting of economic variables and impact analysis. The goal is to improve the results of simply using the most recent table. In this search, however, the gained improvement (if any) must be justified in light of the costs involved.

The choice of technique itself is also subject to constraints imposed in any given situation by the availability of data, personnel, computing abilities and facilities, etc. The methods must also be justified in terms of loss of accuracy (as compared to total survey tables) and the gain in terms of reduced costs (again as compared to total survey tables).

As mentioned previously, there are a large number of suggested methods. A large body of literature that evaluates and analyzes these methods from a variety of points of view. To present the dimensions of this research area, suffice to mention a few major survey papers in this field.

Richardson (1972) offers a valuable, up to date, and comprehensive examination and critique of non-survey techniques. A thorough review of fifty eight projects concerning estimation and projection of I-O coefficients is given by Allen and Gossling (1975). In another excellent survey, Round (1983) evaluates eighty works done by the researchers in this field and brings Richardson's survey up to date. Lastly, Polenske, Crown, and Mohr (1986) present an in-depth review and critique of twenty five research projects dealing with the RAS technique alone.

These researches are primarily descriptive in nature and for the most part an evaluation per se is not offered. However, to the extent that such evaluation exist, the verdict is mixed. Some of the methods consistently perform better than others, while others do not score as well and in several cases the ranking is mixed. In some cases the results are reasonably accurate while in other instances the errors can not be tolerated.

Conceptually, if a relatively large number of survey based tables are available, one can utilize the statistical techniques for trend analysis and extrapolation to update the coefficients. Adequate number of tables for this purpose, however, are rare, and in the limited cases that this approach has been utilized the results are not promising. For instance a study by Tilanus (1966), who used ten tables for the Netherlands economy covering the period 1948-1961, reports results that are inferior to those obtained through application of the most recent survey based table. Barker (1975) did a similar experiment with the data from the U.K for the period 1954-1963 and confirmed these findings. In this case, he also constructed a technology matrix via coefficient by coefficient trend projection. This technology matrix as well as the actual technology matrices for 1954 and 1960 along with the 1963's final demand vector were used to determine the intermediate demand for 1963. These results along with the results obtained through a sector by sector final demand blowup approach were compared and it was concluded that the best result is still obtained via utilization of the most recent survey based table.

Tilanus (1967) used another method for the Netherlands' data for the period 1948-1960. This method, known as "marginal input coefficients" can be described as:

(3-4) $$X^t = X^0 + \Delta X$$

or:

(3-5) $$X^t = [I - A^0]^{-1} Y^0 + [I - A^{0^*}]^{-1} \Delta Y$$

where $\Delta Y = Y^t - Y^0$ or the change in final demand between the two periods, and A^{0^*} is the matrix of marginal input coefficients. The elements of this matrix, a_{ij}^*, are defined as:

(3-6) $$a_{ij}^{*}=\frac{z_{ij}^{0}-z_{ij}^{0-n}}{X_{j}^{0}-X_{j}^{0-n}}=\frac{\Delta z_{ij}}{\Delta X_{j}}$$

Here (0), (t), and (0 - n), represent current, a future and a previous dates respectively. This, as mentioned by Miller and Blair (1985), relates the change in the amount of input (i) purchased by industry (j) for a given period to the change in the total amount of (j) produced over the same period. Since the elements of the intermediate demand matrix for a future date are estimated by utilization

of A^0 and A^{0^*} , in essence change over time of the

coefficients has been entered into the analysis. The reported results, however, were not as good as the results of simply applying the most recent available coefficients.

William Miernyk (1965) Suggests an approach that he calls "best practice firms." The idea is simply to identify the best firm in each industry and the coefficient of the I-O table be determined on the basis of these firms' data. The "best" can be defined using one or several criterions such as profitability, low labor intensity in production, etc. The expectation is that these firms which are currently ahead of the rest of their competitors will be followed in the future. Therefore the most advanced firms of today are the typical firms of tomorrow, hence the composition of their production today will be representative of the structure of production in the future, i.e., they provide relevant technological coefficients of the future.

This method is suggested for near term forecasting and has certain logical appeal and ease of implementation. The problem is the ambiguity inherent in it, that is, unanswered questions like how often these best firms should be identified? How frequently the updating should take place? Why should transition of other firms to the technology of the best firm in their respective industry be uniform and smooth? Why should all industries and within the same time interval make the transition? etc.

Biproportional adjustment or RAS method and its various extensions constitute another set of estimating and updating procedures used in I-O analysis. The biproportional adjustment was first introduced by Deming and Stephen (1940), and Stephen (1942). It was first applied to the I-O analysis by Stone, et al (1963) when they utilized the technique to update the 1960 I-O table of the U.K. They updated the 32 by 32 survey based 1954 table, and used it to project the 1966 one. Subsequently, the RAS method has been expanded and applied extensively. Since the present project will employ the RAS technique and several of its derivatives, a more detailed survey of literature along with presentation of the procedure and its theoretical content will be offered in the next chapter.

The usage of the Lagrangian multiplier as a method of updating and estimating the I-O coefficients has been suggested and utilized by many researchers such as Friedlander (1961), Bacharach (1965), Almon (1968), Geary (1973), and Morrison and Thuman (1980).

The technique in its original form can not guarantee generation of non-negative coefficients and is not capable of accommodating inequality constraints, which sometimes may be needed. In other words, it is possible that a researcher wants to impose some exogenous information on the estimate, e.g., total amount of steel or labor requirement must be less than or equal to certain number or maybe one wants to impose a non-negativity constraint on the model or possibly upper and lower limits on the estimated coefficients, etc.

The Lagrangian method can not accept these constraints in an inequality form and will accept the *a priori* information only at a given level (and not less than or greater than). Thus, for instance, one can either leave the method with no additional constraints and run the risk of having negative estimated coefficients or deviate from a desired level, or set the values equal to a given number which may not be the optimal answer. The Lagrangian method is utilized by the present project, hence a more detailed discussion is deferred to the next chapter.

Another updating technique is a method known as "transaction proportional to value added." This approach basically assumes that all transactions are proportional to value added. That is, the amount of input used by industry (j) from the output of industry (i) is directly proportional to the value added by industry (j). Since this method is used in the current research, a more detailed explanation will be included in chapter four. However, it should be mentioned that application of this technique by Khan (1988) to the data from Pakistan for the period 1975-1985 generated results far inferior to those of RAS. The method overstated total production of Pakistan during the period covered in the study by an average of 16.3%, with average percentage error ranging from 11% to 518%.

A linear programming solution to the problem was devised by Matuszewski, Pitts, and Sawyer (1964) and was used to update the 42 by 42 Canadian I-O table of 1949 to 1956. Their method required a minimal amount of data, was able to adjust a sub-set of the coefficients, and dropped the assumption of uniform change of all coefficients in each industry. They also imposed upper and lower bounds on the value of the estimated coefficients. This method has been used by numerous researchers including Bacharach (1965), Schneider (1965), Davis et al. (1977), and Kim (1984).

Quadratic programming is another technique suggested originally by Omar(1967) and later by Harrigan and Buchanan (1984). This method can overcome problems associated with the Lagrangian method, namely the possibility of generation of negative answers and inability to accommodate inequality constraints. A general statement of this type of problem can be expressed as minimize (3-7)below:

$$\frac{1}{2}\left[\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{(\tilde{x}_{ij}-x_{ij}^{0})^{2}}{x_{ij}^{0}}\right)+\left(\sum_{j=1}^{n}\frac{(\tilde{v}_{j}-v_{j}^{0})^{2}}{v_{j}^{0}}\right)+\left(\sum_{i=1}^{n}\frac{(\tilde{u}_{i}-u_{i}^{0})^{2}}{u_{i}^{0}}\right)+\left(\sum_{t=1}^{n}\frac{(\tilde{C}_{k}-C_{k}^{0})^{2}}{C_{k}^{0}}\right)\right]$$

subject to:

(3-8) $$\underline{u}_i\leq\sum_{j=1}^{n}\tilde{x}_{ij}^{t}\leq\overline{u}_i$$

(3-9) $$\underline{v}_j\leq\sum_{i=1}^{n}\tilde{x}_{ij}^{t}\leq\overline{v}_j$$

(3-10) $$\underline{C}_k\leq\sum_{ij\epsilon k}^{m}\alpha_{(ij)}^{k}\tilde{x}_{ij}\leq\overline{C}_k$$

(3-11) $$\underline{x}_{ij}\leq\tilde{x}_{ij}^{t}\leq\overline{x}_{ij}$$

The solution to the above system can be stated as:

(3-12) $$\tilde{x}_{ij}^{t} = x_{ij}^{0} + \left(\lambda_i + \mu_j + \sum_{k}^{m}\alpha_{ij}^{k}\,\xi_k + \varepsilon_{ij} - \kappa_{ij}\right)x_{ij}^{0}$$

where $\tilde{x}_{ij}^{t}$'s are the estimated elements of the intermediate

demand and x_{ij}^{0} 's are their corresponding elements in the base year table. (k) represents additional constraints imposed on elements of the updated matrix. u_i^t , v_j^t , c_k^t are elements of the updated vectors of column sums, row sums, and imposed constraints respectively, and, u_i^0 , v_j^0 , c_k^0 indicate their corresponding initial best estimates.

Here, unlike the Lagrangian method constraints are treated as variables and not constants. The upper and lower bars indicate exogenously set limits of variables, and the $\alpha_{ij,k}^t$,s are constant coefficients providing for the expression of proportionality relationships between individual and aggregates of elements. λ_i , μ_j , ξ_k are the Lagrangian multipliers, and ε_{ij} , κ_{ij} denote the upper and lower limits imposed on the estimate of each element of the updated matrix.

More explanation of this technique and its application to the Korean tables for 1970, 1975, and 1980 can be found in Kim (1984). According to him, the results obtained through the method are operationally acceptable and only marginally different from those obtained via RAS method. A detailed solution of quadratic programming problem, along with its application to 1963-1967 data for the state of Washington, a well as the comparison of the results with those of the RAS method is also presented in Harrigan and Buchanan (1984).

A technique suggested by the Economic Planning Agency in Japan, called the "Residual Minimum Method" is yet another updating method. Fundamentally, this method is the same as the Lagrangian technique with the exception of introduction of an adjusting weight factor into the objective function. Since this approach has been used in the present project, a detailed description will be given in the next chapter. Kanko(1983) who applied this method to Japanese tables for 1970-1977 reports a coefficient of equality and the standard deviation (.925 and .287) that are compatible with those of RAS (.928 and .288). On cell by cell basis this method estimated 252, 387, and 470 of the cells that fall within plus or minus 5%, 10%, and 20% of the actual values respectively. The reported numbers for the RAS procedure are 154, 279, and 362.

In addition to the aforementioned techniques there are some estimation methods that are devised exclusively for regional tables, i.e., they address the question of "spatial stability" of the coefficients. The need for these procedures arises from the fact that due to various reasons (time, budget, personnel, etc.) regional economists do not have a survey based table pertaining to their particular region. Hence, they attempt to construct the desired tables from the national tables, i.e., "regionalizing" them.

Regionalization process basically involves estimation of regional technical coefficients from their national counterparts. The question, however, is the applicability and reliability of these practices. To what extent the national coefficients remain stable at the regional level and how much confidence can be placed on the results obtained through regeonalized tables?

There is an enormous body of literature on the question of regionalization and its subsequent issues. Numerous researchers have tried to answer these questions from a variety of perspectives and for different circumstances. The results, among other things, have been many techniques and procedures of estimating and updating the I-O coefficients. Notable among these techniques are, procedures such as: "purchase only location quotient," "sales only quotient," "simple location quotient," "cross industry location quotient," "modified cross industry location quotient," "logarithmic cross industry location quotient," "modified logarithmic cross industry location quotient," "fabrication effects," "regional purchase coefficients," "supply-demand pool," "commodity balance technique," "regional Input-Output simulator," "entropy methods," and host of other such methods. However, since the present study is concerned with the national data, hence temporal stability of the coefficients, the question of spatial stability falls outside of its domain. Therefore, the review and analysis of this category of estimating and updating methods will not be pursued here. A rather comprehensive survey of these techniques can be found in Sawyer and Miller (1983), Miller and Blair (1985), and Polenske et al. (1986).

The search for more techniques and various circumstances to which they are applied or can be applied may continue substantially more than what was presented in this chapter. In all likelihood, however, very little additional insight will be gained in this path. All major studies on the subject have been examined and the crux of the matter is presented. Out of this search, emerged a list of techniques for the purpose of the empirical part of the current research. These techniques are selected through consideration of several factors such as applicability, plausibility, efficiency, past performance, acceptability in the field, etc. In the interest of clarity and comprehension, however, before entering the empirical domain of the current study and verification of the merits of these methods, a rather detailed description of foundation, mechanism, and structure of the chosen updating techniques will be presented in the next chapter.

CHAPTER FOUR

THE ESTIMATION AND UPDATING TECHNIQUES

"I don't rejoice in insects at all," Alice explained;
"because I am rather afraid of them - at least the large kinds. But I can tell you the names of some of them."
"Of course they answer to their names?" the Gnat remarked carelessly.
"I never knew them to do it."
"What is the use of their having names," the Gnat said,
"if they won't answer to them?"
"No use to them," said Alice; "but it's useful to people that name them, I suppose. If not why do things have names at all?"

Alice in Wonderland and Through the Looking Glass

Review of the literature led to selection of three estimation techniques for utilization in this project. These methods are RAS, Lagrangian, and "Proportional to Value Added." The choice, as mentioned earlier, is based on several considerations. Chief among them are wide spread usage, acceptability in the field, logical plausibility, computational ease, and minimal information requirement.

These basic methods, along with some of their variations, extensions, and combinations provide eight estimation techniques to be applied to the data. The NAIVE, or constant coefficient case is also added to the list for both comparative purposes as well as a test of the assumption of coefficients' temporal stability.

Before presenting the results of application of these nine methods and conducting comparisons, however, some detailed explanation of the methods is in order. The current chapter is an attempt to probe the theoretical and structural foundations of the selected techniques, which, in turn, sets the stage ready for evaluation, comparison, and explanation of the empirical results.

4.1) THE RAS PROCEDURE:

The RAS procedure is the most widely used technique for estimation of I-O coefficients. This popularity is owed to its relative modest informational requirements and the relative availability of the data that it requires. This, in turn, reduces to a minimum the time and expenditures needed for updating of an I-O table. Adding to this factor the relative accuracy and power of the RAS procedure, as indicated by numerous studies, along with its ability to preserve signs and zero values of the coefficients, makes the technique a very attractive one.

For these reasons, the present study utilizes the RAS technique as well as number of its variants. Therefore, somewhat detailed presentation of this technique along with various surrounding issues seems appropriate.

As noted earlier, biproportional matrix adjustment was first suggested by Deming and Stephen (1940), and Stephen(1942). They employed an iterative procedure to adjust the 1940 U.S census population data. The first mention of biproportional adjustment in order to take account of changes in I-O coefficients was made by Leontief (1941), and its first comprehensive application to the I-O field was by Stone (1961), followed by Stone and Brown (1962), and Stone et.al. (1963) through Cambridge University's *"A Program for Growth"* series.

In these studies the 1960 I-O table of the U.K. was estimated, using the 1954 table as the benchmark data, along with row and column totals of 1960. Bacharach (1970) discussed the mathematical properties of RAS and showed that the procedure basically involves minimization of an objective function such as $\sum_{i}^{n}\sum_{j}^{n}\left[\tilde{x}_{ij}^{t}\ln\left(\frac{\tilde{x}_{ij}^{t}}{x_{ij}^{0}}\right)\right]$ subject to the row and column marginal constraints. He also provided proof of the convergence, uniqueness of the solution, and preservation of zero elements. Further proof of these properties of RAS and its extension to the case of "modified RAS" was furnished by MacGill (1977 and 1979).

Ever since its introduction, much research has been conducted on various aspects, applications , and extensions of the RAS procedure. Almost all researchers dealing with comparative efficiency of updating methods have included RAS in their project as one of the methods. In fact, in most cases the RAS procedure serves as the "norm," against which the efficiency of other methods is measured.

In addition to the works previously cited, many other researches in this area can be mentioned. Early works on RAS include Schneider (1965) on the 1947-1958 data for the U.S., Paelinck and Waelbroeck (1965, cited in Bacharach, 1970) on the 1953-1959 Belgian data, Grandville, et al. (1968) for projection of foreign trade, Czamanski and Malizia (1969) for construction of the Washington state's table of 1963 based on 1958 U.S. table, Schaffer and Chu (1969) to obtain the 1963 Washington state's table from 1958 table of the U.S., Henry (1973) and (1974) on estimation of Irish table, Morrison and Smith (1974) to estimate 1968 table for the Peterborough, based on 1968 U.K table, McMenamin and Haring (1974) on data for 1963 Washington state table to estimate 1967 one, Malizia and Bond (1974) for estimating 1967 table of Washington state based on 1963 table, Lamel et al. (1974) for estimating 1968 Norwegian table from the 1963 table, Allen and Lecomber (1975) for the U.K data in the period 1954-1968, and Miernyk (1975) for estimation of the U.S 1967 table based on the 1963 table. Other notable early experiments with RAS includes Omar (1967), Lecomber (1969), Allen (1974), and Barker (1975 A) on the U.K data.

Later experiments with RAS encompasses studies by Hewing (1977) for adjusting non-local coefficients, i.e., utilization of tables from Washington and Kansas states in estimation of one table via utilization of the other, Hinojosa (1978) on the 1963, 1967, and 1972 data for the state of Washington, Parikh (1979) for the estimation of the 1965 coefficients of nine European countries based on the U.N. data, Harrigan, et.al., (1980) for the estimation of 1973 I-O table of Scotland based on 1973 table of the U.K., Hewings and Syverson (1980, cited in Polenske et al. 1986) on construction of the tables for Washington state based on 1963, 1967, and 1972 tables, and Butterfield and Mules (1980) in estimation of 1958-1959 regional table of Western Australia from the national table of the same period.

More recent application, evaluation and extension of the RAS method includes researches by Sawyer and Miller (1983) on construction of 1972 table for the state of Washington based on the US 1967 table, Kim (1984) on accuracy comparison using Korean data for the 1970-1980 period, Pigozzi and Hinojosa (1985) on modifying the method, Cray (1986) in updating the 1969 table for the state of

Kansas to 1982, Israilevich (1986) on extension of the method and application of this extension to the 1963-1967 data for the U.S., Lynch (1986) on estimation of 1968 table of U.K from the 1963 table, Khan (1988) on updating of the I-O table of Pakistan from 1975/76 to 1984/85, St. Louis (1989) on experimentation with Canadian data, and Snower (1990) on extension of the technique.

The findings of these studies, as well as those of previously mentioned researches can broadly be summarized. The I-O coefficients demonstrate a surprising degree of stability in the short term. With the passage of time, however, the coefficients become unstable and the application of latest ("old") survey based techniques for the medium and long term purposes will not produce reliable results. The RAS method (and its variants), in these cases usually performed better than simply applying the most recent coefficients. They also either outperformed or matched all other non-survey techniques of estimating and updating I-O coefficients. The findings also indicate that although the overall accuracy of RAS (and its variants) in generating aggregate estimates is reliable and operationally acceptable for economic analysis (for up to fifteen years in advance as reported by Grandville, et. al. 1968), but in many cases it produces poor and unreliable, and sometimes unacceptable, estimates for individual cells in the table. At the sectoral level also the accuracy of the estimates are not uniform. While in many cases poor results were generated for some sectors such as service trade, transportation and communication, manufacturing sectors are generally reasonably estimated. These researches also seem to indicate that estimates can be improved via incorporation of additional data on certain sectors as well as inclusion of other information (e.g. structural changes, policy changes, technological changes, etc.) into the updating procedure. This latter issue will be discussed in more detail in chapter seven.

4.2) FORMAL PRESENTATION OF RAS TECHNIQUE:

The information requirement of the RAS procedure, as was noted earlier, is relatively modest. Specifically, one needs only one complete survey based interindustry matrix along with row and column totals for a later date interindustry matrix to be able to employ RAS for estimation of the elements of the second matrix. This is to say that to estimate an (n) by (n) technology matrix of a future date, A^t , in addition to a benchmark matrix of present or past date, A^0 , one only needs the vector of total gross outputs for the n industries in the updating year (X_j^t) , the vector of total intermediate sales by each of the n industries $\left(\sum_{j=1}^{n} x_{ij}^{t}\right)$, and the vector of total intermediate purchases by each of the n sectors $\left(\sum_{i=1}^{n} x_{ij}^{t}\right)$. This simply implies that in order to obtain n^2 pieces of information about the future (i.e. the estimates of a_{ij}^{t} 's) only 3n pieces of data with regards to that date are needed, which in turn translates into substantial savings of effort and expenditure.

The technique itself revolves around the assumption that the effects of all factors causing change in the coefficients of the benchmark table, from the base year to the target year, can be captured by biproportional relationships where each coefficient is subject to a row wise and column wise adjusting multiplier. These multipliers, denoted by r_i and s_j respectively are further assumed to uniformly affect the rows and columns of the matrix. The problem, then, is to obtain values for r_i

and s_j in such a way that when applied to the base year

matrix, generates a new matrix whose elements satisfy the predetermined row and column marginals. Formally the method can be presented as follows:

(4-1) $$\tilde{X}^t = \hat{R}\, X^0\, \hat{S}$$

subject to:

(4-2) $$\tilde{X}^t i = \sum_{j=1}^{n} r_i x_{ij}^0 s_j = u_i^t$$

and:

(4-3) $$i'\tilde{X}^t = \sum_{i=1}^{n} r_i x_{ij}^0 s_j = v_j^t$$

Or alternatively

(4-4) $$\tilde{A}^t = \hat{R}\, A^0\, \hat{S}$$

subject to:

(4-5) $$(\tilde{A}^t \hat{X}_j^t)\, i = U_i^t$$

and:

(4-6) $$i' (\tilde{A}^t \hat{X}_j^t) = V_j^t$$

where $\hat{R}$ and $\hat{S}$ are diagonal matrices constructed from the vectors of row and column wise multipliers (r_i 's and s_j ,s) respectively.

$\tilde{A}^t$ = estimated matrix of coefficients for target year (t).

A_0 = the base year's matrix of coefficients.

x_{ij}^0 = element of intermediate demand for i-th commodity in j-th industry in the base year.

U_i^t = the total intermediate output vector for the year(t).

V_j^t = the total intermediate input vector for the year (t).

u_i^t and v_j^t = elements of U_i^t and V_j^t respectively.

i' and i = unit row and column vectors respectively.

$\hat{X}^t$ = diagonalized matrix of total gross output for the period (t).

$\tilde{X}^t$ = the target year's estimated intermediate transaction matrix.

X^0 = the base year's intermediate transaction matrix.

This procedure involves successive biproportional adjustment of the rows and columns of the base year matrix until convergence is achieved. The process can be explained as follows:

Assume that the row and column sums of the transaction matrix X^t for the target year (i.e., U_i^t and V_j^t) are known.

Further assume that the vector of total gross output for the target year along with the base year's technology matrix are also available. In other words the followings are known:

(4-7) $$[A^0] = \begin{bmatrix} a_{11}^0 & \cdots\cdots & a_{1n}^0 \\ \cdot & \cdots\cdots & \cdot \\ \cdot & \cdots\cdots & \cdot \\ a_{n1}^0 & \cdots\cdots & a_{nn}^0 \end{bmatrix}$$

and $[V_j^t] = [v_1 \; . \; . \; . \; . \; v_n]$, as well as:

$$[X_j^t] = \begin{bmatrix} X_1^t \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ X_n^t \end{bmatrix} \qquad \text{and} \qquad [U_i^t] = \begin{bmatrix} U_1^t \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ U_n^t \end{bmatrix}$$

Constancy of the coefficients between the two time periods implies $A^t = A^0$. Since $X^t = A^t \hat{X}_j^t$,then if $A^0 = A^t$,

utilizing the known values for the target year, it is evident that:

(4-8) $$[A^0 \hat{X}_j^t] = \begin{bmatrix} a_{11}^0 X_1^t, & a_{12}^0 X_2^t, & . \; . \; . \; ., & a_{1n}^0 X_n^t \\ a_{21}^0 X_1^t, & a_{22}^0 X_2^t, & . \; . \; . \; ., & a_{2n}^0 X_n^t \\ \cdot & \cdot & \cdot \; \cdot \; \cdot & \cdot \\ \cdot & \cdot & \cdot \; \cdot \; \cdot & \cdot \\ \cdot & \cdot & \cdot \; \cdot \; \cdot & \cdot \\ a_{n1}^0 X_1^t, & \cdot & . \; . \; . \; ., & a_{nn}^0 X_n^t \end{bmatrix}$$

Postmultiplying (4-8) by a unit column vector and denoting the result as U_1^t will yield:

(4-9) $U_1^t = [A^0 \hat{X}_j^t][i]$

where U_1^t can be considered as the first estimate of U_i^t . If these two are equal, then it can be concluded that the estimate of the target year's interindustry transaction (i.e., $\hat{X}_j^t$)is accurate with respect to the row sums, hence only the column sums must be verified. This test can be conducted through premultiplication of $[A^0 \hat{X}_j^t]$ by a unit row vector and comparing the results with V_j^t . That is to say whether $V_j^1 = [i'][A^0 \hat{X}_j^t]$ or not. If these conditions hold, then the estimated transaction matrix must be identical to the actual transaction matrix of the target year. Consequently, the matrices of technological coefficients for the base and the target year will be equal, viz, intertemporal stability of the coefficients. However, if these equalities do not hold, i.e., $U_i^1 \neq U_i^t$ or $V_j^1 \neq V_j^t$, which is the more likely situation, some additional adjustments are in order. Violation of these equalities, for instance

$U_i^1 < U_i^t$,indicates that the estimates of the elements of i-th row are smaller than they should be or $V_j^1 > V_j^t$ indicates the overestimation of the elements of the j-th column. Other possibilities can be interpreted similarly. The adjustment mechanism used in RAS operates as follows:

Let $\frac{u_i^t}{u_i^1} = r_i^1$. This ratio, depending on the relative values of numerator and denominator, can be less than or greater than one. Therefore multiplication of each element of the i-th row in A^0 by r_i^1 will reduce (increase) their values. These new values (new coefficients), when multiplied by their corresponding elements in the interindustry matrix, still satisfy the row sums constraint for the i-th sector. Hence they are improved estimates of the target year's coefficients in the i-th sector. The process continues for all rows of the base year matrix until a new estimate of the matrix of coefficients for the target year is generated. The product of this matrix (i.e., A^1) and the target year's interindustry sales will satisfy the known row sums vector of total interindustry demand for the target year. Algebraically,

$$\text{(4-10)} \qquad [A^1] = \begin{bmatrix} r_1^1 & 0 & 0 & . & . & . & 0 \\ 0 & r_2^1 & 0 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & . & r_n^1 \end{bmatrix} [A^0]$$

Letting the term in the first bracket to be R^1 will yield:

$$\text{(4-11)} \qquad A^1 = R^1 A^0$$

where $R^1 = [\hat{U}_i^t]\,[\hat{U}_i^1]^{-1}$, since elements of R^1 are the ratio of their corresponding elements in U_i^t and U_i^1 . These adjustments will ensure the constraints implied by U_i^t are satisfied, since the violation of these constraints was the motif behind the adjustment in the firs place. Therefore:

$$\text{(4-12)} \qquad [R^1\, A^0\, \hat{X}_j^t]\,[i] = [A^1\, \hat{X}_j^t]\,[i] = X^1[i] = U_i^t$$

Having the row sums conditions met, the issue will turn to verification of the column sums constraints. That is to see if $V_j^t = [i]'[A^1\hat{X}_j^t]$. If this equality holds, then A^1 is the desired updated matrix. If, however, the equality is violated (i.e., if $V_j^1 \neq V_j^t$), a second set of modification must be performed. Here too, $V_j^1 > V_j^t$ indicates that the column sums of A^1 are larger than they should be and visa versa.

The adjustment starts by setting $s_j^1 = \frac{V_j^1}{V_j^t}$.

Postmultiplying each column of A^1 by its respective s_j^1 will yield:

$$\text{(4-13)} \qquad [A^2] = [A^1] \begin{bmatrix} s_1^1 & 0 & . & . & . & . & 0 \\ 0 & s_2^1 & 0 & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & s_n^1 \end{bmatrix}$$

which gives the second estimate of A^t ,that is A^2 .

Denoting the term in the bracket by S^1 , one can get:

(4-14) $$A^2 = A^1 S^1$$

It is evident that $S^1 = [\hat{V}_j^t][\hat{V}_j^1]^{-1}$. The second round of adjustment ensures that the column sums restrictions are met,i.e., $X^2 = A^2[\hat{X}_j^t]$ and $[i]'[A^2 \hat{X}_j^t] = [i]'[X^2] = V_j^t$. At this point A^2 must be checked for row sums constraints. If these are met, then the final estimate of the matrix for the target year is at hand. If that is not the case, another round of modification must be performed. That is to say, if $[A^2 \hat{X}_j^1][i] = U_j^2 \neq U_j^t$, the process of adjustment must continue via construction of a row-modifying matrix. Let $r_i^2 = \frac{u_i^t}{u_i^2}$ and define $\hat{R}^2$ as a diagonal matrix formed by these elements. Thus:

$$(4\text{-}15)\qquad [R^2] = \begin{bmatrix} r_1^2 & 0 & 0 & . & . & . & 0 \\ 0 & r_2^2 & 0 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & . & r_n^2 \end{bmatrix} = [U_i^t]\,[U_i^2]^{-1}$$

If the value obtained from this equation is equal to $[I]$ then all row and column constraints are satisfied and the problem is solved. If that is not the case, further adjustment is warranted. This suggests that $A^3 = R^2A^2$ must be constructed. This equation, again, guarantees the satisfaction of row marginals.

The column marginals, however, must be verified. Thus,the $[i']\,[A^3]\,[\hat{X}_j^t] = V_j^2$ will be checked against V_j^t . If they are equal A^3 is the answer, if not another matrix of column-modifier should be constructed. That is to say:

$$(4\text{-}16)\qquad A^4 = [A^3]\,[S^2]$$

where S^2 is the diagonal matrix of modifiers constructed from $s_i^2 = \dfrac{v_i^t}{v_i^2}$ or simply $S^2 = [\hat{V}_j^t]\,[\hat{V}_j^2]^{-1}$. Then the next.

estimate is generated via $A^4 = A^3S^2$. Through comparison of equations (4-11) and (4-14), it is clear that $A^2 = R^1A^0S^1$. By the same logic:

(4-17) $$A^3 = [R^2]\,[R^1]\,[A^0]\,[S^1]$$

or:

(4-18) $$A^3 = [R^2R^1]\,[A^0]\,[S^1]$$

Utilizing equations (4-16) and (4-18) will yield:

(4-19) $$A^4 = [R^2R^1]\,[A^0]\,[S^1]\,[S^2]$$

or:

(4-20) $$A^4 = [R^2R^1[A^0]\,[S^1S^2]$$

R^2R^1 and S^1S^2 in this equation are (n) by (n) matrices constructed from multiplication of R^2 by R^1 and S^2 by S^1 respectively. Therefore:

(4-21) $$[R^2R^1] = \begin{bmatrix} r_1^2r_1^1 & 0 & \dots & 0 \\ 0 & r_2^2r_2^1 & 0 \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & r_n^2r_n^1 \end{bmatrix}$$

and:

(4-22) $$[S^1S^2] = \begin{bmatrix} s_1^1s_1^2 & 0 & 0 \dots & 0 \\ 0 & s_2^1s_2^2 & 0 \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & s_n^1s_n^2 \end{bmatrix}$$

Continuing the process will yield the sequence (4-23) below:

$$A^5 = [R^3R^2R^1]\,[A^0]\,[S^1S^2]$$
$$A^6 = [R^3R^2R^1]\,[A^0]\,[S^1S^2S^3]$$
$$\dots\dots\dots$$
$$A^{(2n)} = [R^nR^{n-1}\dots R^1]\,[A^0]\,[S^1S^2\dots\dots S^n]$$

Denoting the terms in the first and the last brackets by R and S respectively, the last equation can be written as:

(4-24) $$A^{2n} = [\hat{R}]\,[A^0]\,[\hat{S}]$$

The procedure will continue until convergence occurs or a predetermined level of closeness is achieved, that is to say, until all elements of the U and V (as computed by

successive rounds) are within a given percentage (say $\varepsilon=5\%$)

of their corresponding elements in U_i^t and V_j^t . Then:

(4-25) $$\tilde{A}^t = [\hat{R}]\,[A^0]\,[\hat{S}]$$

where $\tilde{A}^t$ is an acceptable estimate of A^t , hence signifies

the conclusion of adjusting procedure.

Mathematical properties of this technique was outlined in Stone, et.al. (1963). Rigorous proof of its properties was given first by Bacharach (1970) and later by several others, e.g., Macgill (1977) Snickars and Weibull (1977), Macgill (1979), and Israilevich (1986). In general, it can be shown that:

a) A unique solution to the problem does exist and that convergence will be reached. This implies that after each round of adjustment of rows and columns, the obtained values are closer to the true values of U^t and V^t .

b) The procedure will preserve all signs. This is ensured due the fact that all row and column modifiers, i.e.,

r_i 's and s_j 's, are positive numbers by their very definition. Therefore the RAS procedure does not produce negative coefficients (a_{ij} 's).

c) All zeros in the original (base year) matrix will be preserved. This is because when these zeros are multiplied by the modifiers, will render the products zero, hence maintaining zeros in the updated matrix corresponding to the zero values in the original one.

d) From an information point of view, least biased results are generated by the RAS procedure.

4.3) THEORETICAL FOUNDATION OF THE RAS METHOD:

All non-survey techniques of estimating and updating of I-O coefficients attempt to accomplish the task in such a way that the estimated table be as close to the original survey based table as possible. This "closeness," however, can be defined and measured in variety of ways. A linear measure is suggested by Matuszewski, Pitts, and Sawyer (1964), i.e., the estimation via linear programming. A least square approach, that is utilization of Lagrangian optimization technique for measuring this distance, is proposed by Morrison and Thumann (1980). And usage of "information theory" for measurement of this closeness is suggested by Bacharach (1970). The latter measurement is what is used by RAS and it is based on the "information theory" proposed by Kullback and Leibler (1951), and as noted by Bacharach (1970) is the only one with a theoretical-probabilistic explanation.

4.3.1) **Information Theory and RAS**. The usage of information theory in economic forecasting was first utilized, in an indirect way, by Uribe, de Leeuw, and Theil (1965). The theory and its economic applications was detailed later by Theil (1967) and Bacharach (1970). Since the understanding of this theory is crucial for understanding of the theoretical foundation of RAS procedure, a brief explanation is offered. The description here, makes use of Kullback and Leibler (1951), Theil (1965), Bacharach (1970), Miller and Blair (1985), and Israilevich (1986). More detailed presentations may be found in these sources.

The notion of "information content" is the foundation of information theory. The information content of an event (E) with probability of(π) is designated as $I(\pi)$ and is measured as:

$$I(\pi) = \ln\left(\frac{1}{\pi}\right) \tag{4-26}$$

which by a simple logarithmic theorem can be rewritten as:

$$I(\pi) = \ln\left(\frac{1}{\pi}\right) = -\ln\pi \tag{4-27}$$

It is evident here that the higher the probability of an event, the smaller the amount of information obtained.

Assume that a researcher, using available data and a simple model, predicts an increase in the Dow-Jones industrial average for a given day with unconditional probability of(π_0). Further assume that a research institute, utilizing a sophisticated model and detailed information, predicts the rise in the stock market for the same day with the probability of(π_1). The difference between these two predictions, termed the "information gain," is the measure of the value or the index of improvement of the second prediction and can be shown as:

$$\iota(\pi_0) - \iota(\pi_1) = \ln\left(\frac{1}{\pi_0}\right) - \ln\left(\frac{1}{\pi_1}\right) \tag{4-28}$$

The information gain relationship above, via employment of simple logarithmic theorems, can be rewritten as:

$$(4\text{-}29)\quad \iota(\pi_0)-\iota(\pi_1) = -\ln\pi_0 - (-\ln\pi_1) = \ln\pi_1 - \ln\pi_0 = \ln\left(\frac{\pi_1}{\pi_0}\right)$$

In case that the $\pi_0 = \pi_1$, the information gain obtained via the message delivered by the research institution is zero.

Now, suppose the researcher wants to assign probabilities to the three possible outcomes, i.e., the index will go up, goes down or stays the same. Denoting these three events by(E_1),(E_2), and(E_3) respectively with their corresponding probabilities shown as (π_1), (π_2), and(π_3), it is clear that for these complementary events:

$$(4\text{-}30)\qquad \sum_{i=1}^{3} \pi_i = 1$$

and the "expected information content" for these events is:

$$(4\text{-}31) \qquad I(\pi) = \sum_{i=1}^{3} \pi_i \, \iota(\pi_i) = \sum_{i=1}^{3} \ln\left(\frac{1}{\pi_i}\right) = \sum_{i=1}^{3} -\pi_i \ln(\pi_i)$$

In the absence of any information about the events and their distribution, the researcher has to assign equal probability to each of these events, i.e., $\pi_1 = \pi_2 = \pi_3$. This is so because optimizing equation (4-31) subject to (4-30) will result in the Lagrangian multiplier to be $\lambda = -1 - \ln(\pi_i)$.

If, however, the researcher obtains the forecasting institution's prediction about these events and he/she generally trusts the forecasts by this particular institution and wants to make his/her own prediction, the situation is different.

Suppose the researcher has access to some additional information that he/she believes was not available to the research institute, hence not taken into the consideration by them in their forecasts. These "signals," alter the researcher's knowledge of the events and change the respective likelihood of occurrence of each event. Under the influence of these signals, assume that the researcher believes that $\pi_1 > \pi_2$ and $\pi_1 > \pi_3$. Let the new set of probabilities assigned by him/her to the events be denoted as: ϖ_1 , ϖ_2 , and ϖ_3 respectively. These are called "posterior" probabilities, while the original probabilities given by the institution are known as "prior" probabilities. The question is how should he/she determine the posterior probabilities, given the fact that on one hand he/she trusts the institution's forecast, but on the other hand has received new signals that he/she is certain were not available to the institution?

Based on information theory, the researcher will attempt to minimize the amount of "surprise" or "information inaccuracy of prediction," i.e.,

$$\text{(4-32)} \quad Min.\ I(\varpi:\pi) = \varpi_i[\iota(\varpi_i) - \iota(\pi_i)] = \varpi_i \ln\left(\frac{\pi_i}{\varpi_i}\right)$$

which implies minimization of the information difference weighted by the posterior probabilities of the events. Moreover, the additional constraints available to him/her must be imposed on the objective function, namely: $\pi_1 > \pi_2$ and $\pi_1 > \pi_3$. It should be noted that this information gain is always greater than zero except when $\pi = \varpi$ which indicates no information gain.

The distribution of transaction flows of an I-O table can be viewed in much the same way. The elements of the interindustry flow matrix, however, do not represent probabilities as such. To overcome this difficulty, they must be "normalized." The normalization can be accomplished via division of each element of the matrix by total of the flows in their respective columns, that is to find relative weight of each cell in each industry. Then the base year matrix will provide the prior probabilities and the information about the new matrix serve as signals for determining the posterior probabilities. Formally:

$$\text{(4-33)} \qquad I\left(\frac{X^t}{X_j^t} : \frac{X^0}{X_j^0}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left[\frac{x_{ij}^t}{X_j^t} \ln \frac{\dfrac{x_{ij}^t}{X_j^t}}{\dfrac{x_{ij}^0}{X_j^0}}\right]$$

or:

$$\text{(4-34)} \qquad I\left(\frac{X^t}{X_J^t} : \frac{X^0}{X_j^0}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{x_{ij}^t}{X_j^t}\left(\ln\frac{x_{ij}^t}{X_j^t} - \frac{x_{ij}^0}{X_j^0}\right)$$

which can be rewritten as:

$$I\left(\frac{X^t}{X_j^t} : \frac{X^0}{X_j^0}\right) = \frac{1}{X_j^t}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(x_{ij}^t\, lnx_{ij}^t - lnX_j^t - lnx_{ij}^0 + lnX_j^0)\right]$$

(4-35)

and rearranged as:

(4-36) $$I\left(\frac{X^t}{X_j^t}:\frac{X^0}{X_j^0}\right) = \frac{1}{X_j^t}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}^t \ln\frac{x_{ij}^t}{x_{ij}^0} + \ln\frac{X_j^0}{X_j^t}\right]$$

Since the term outside the bracket is a scaler and the last term inside the bracket is a constant, the solution will be identical to the solution prior to normalization. This means the base year matrix and additional information about the target year can be utilized directly to arrive at this solution via information theory. Information theory implies that the posterior probabilities assigned by this objective function have the lowest information content. According to Snickars and Weibull (1977), this is the only way to estimate transaction flows via information theory.

Equation (4-36) and the its related imposed constraints can be written in terms of I-O coefficients as the minimization of the information distance between

A^0 and A^t, i.e.,

(4-37) $$I(A^0:A^t) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^t \ln\left(\frac{a_{ij}^t}{a_{ij}^0}\right)$$

Subject to:

(4-38) $$\sum_{j=1}^{n} a_{ij}^t X_j^t = U_i^t$$

and:

(4-39) $$\sum_{i=1}^{n} a_{ij}^{t} X_j^t = V_j^t$$

Setting the problem in the Lagrangian format will yield:

$$L = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^{t} \ln\left(\frac{a_{ij}^{t}}{a_{ij}^{0}}\right) - \sum_{i=1}^{n} \lambda_i \left[\sum_{j=1}^{n} a_{ij}^{t} X_j^t - U_i^t\right] - \sum_{j=1}^{n} \mu_j \left[\sum_{i=1}^{n} a_{ij}^{t} X_j^t - V_j^t\right]$$

(4-40)

Taking partial derivative with respect to a_{ij}^{t} and setting it

equal to zero, will result:

(4-41) $$\frac{\partial L}{\partial a_{ij}^{t}} = 1 - \ln(a_{ij}^{t}) - \ln(a_{ij}^{0}) - \lambda_i X_j^t - \mu_j X_j^t = 0$$

or:

(4-42) $$\ln(a_{ij}^{t}) = \ln(a_{ij}^{0}) - 1 + \lambda_i X_j^t + \mu_j X_j^t$$

Taking antilogarithm will give:

$$a_{ij}^{t} = [e]^{(\ln a_{ij}^{0} - 1 + \lambda_i X_j^t + \mu_j X_j^t)} = [e]^{(\ln a_{ij}^{0})}[e]^{(-1 + \lambda_i X_j^t + \mu_j X_j^t)}$$

(4-43)

or:

(4-44) $$a_{ij}^{t} = a_{ij}^{0}\ [e]^{(-1 + \lambda_i X_j^t + \mu_j X_j^t)}$$

which by rearranging can be rewritten as (4-45) below:-

$$a_{ij}^{t} = [a_{ij}^{0}]\,[e]^{(\lambda_i X_j^t - \frac{1}{2})}\,[e]^{(\mu_j X_j^t - \frac{1}{2})} = [e]^{(\lambda_i X_j^t - \frac{1}{2})}\,[a_{ij}^{0}]\,[e]^{(\mu_j X_j^t - \frac{1}{2})}$$

denoting the first and second brackets as r_i and s_j , then:

(4-46) $$a_{ij}^{t} = [r_i]\,[a_{ij}^{0}]\,[s_j]$$

It is evident that the first bracket is a function of λ_i only which means it is a row constraint, and the third bracket is a column constraint, hence a function of μ_j only. This in turn, can be seen as the adjustment of the base year's coefficients by row and column constraints terms r_i and s_j to obtain the target year's coefficients.

The values for r_i 's and s_j 's are derived through taking the appropriate partial derivatives and equating them to zero. That is:

(4-47) $$\frac{\partial L}{\partial \lambda_i} = \sum_{j=1}^{n} a_{ij}^{t} X_j^{t} - U_i^{t}$$

and:

(4-48) $$\frac{\partial L}{\partial \mu_j} = \sum_{i=1}^{n} a_{ij}^t X_j^t - V_j^t$$

Substituting for a_{ij}^t from equation (4-46), in (4-47) and (4-48), and equating them to zero will give:

(4-49) $$\sum_{j=1}^{n} [r_i] [a_{ij}^0] [s_j] [X_j^t - U_i^t] = 0$$

(4-50) $$\sum_{i=1}^{n} [r_i] [a_{ij}^0] [s_j] [X_j^t] - V_j^t = 0$$

or:

(4-51) $$r_i = \frac{U_i^t}{\sum_{j=1}^{n} [a_{ij}^0] [s_j] [X_j^t]}$$

and:

(4-52) $$s_j = \frac{V_j^t}{\sum_{i=1}^{n} [r_i] [a_{ij}^0] [X_j^t]}$$

An iterative solution to (4-51) and (4-52) will yield the desired values of r_i and s_j . The resultant a_{ij}^t can be verified to be indeed the minimum value by taking the second derivative in equation (4-40) and setting it equal to zero. That is:

(4-53) $$\frac{\partial^2 L}{\partial^2 a_{ij}^t} = \frac{1}{a_{ij}^0}$$

Since all a_{ij}^0 's are positive, the value of equation (4-53) is positive, hence indicating a minimum point.

For the entire matrix of the target year, A^t , two vectors of r_i ,s and s_j 's are provided via RAS technique. Diagonalizing these two vectors will yield:

(4-54) $$[R] = \begin{bmatrix} r_1 & 0 & 0 & . & . & . & 0 \\ 0 & r_2 & 0 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & . & r_n \end{bmatrix}$$

and:

(4-55) $$[S] = \begin{bmatrix} s_1 & 0 & 0 & . & . & . & 0 \\ 0 & s_2 & 0 & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & s_n \end{bmatrix}$$

or generally:

(4-56) $$A^t = [\hat{R}]\,[A^0]\,[\hat{S}]$$

which is the same as equation (4-1). It is evident that premultiplication of A^0 by the diagonal matrix $\hat{R}$, uniformly adjusts each element of the rows of the base year's matrix. Similarly, postmultiplication of A^0 by diagonal matrix $\hat{S}$, uniformly adjusts each element of its columns.

4.3.2) **Economic Basis Of RAS**. As noted earlier, in the RAS procedure elements of the base matrix are adjusted by $\hat{R}$ and $\hat{S}$, which are products of diagonal matrices R^n.......R^1 and S^1......S^n respectively. Each element of $\hat{R}$ (i.e., r_i) and $\hat{S}$ (i.e. s_j) is multiplied by each element of i-th row and j-th column of the base year matrix respectively for i and j =1,2,....,n.

The problem, however, is that despite the relative power of this procedure in updating the base year's matrix - as it is evident from the empirical tests - there is no clear economic justification of these proportional uniform changes. The most common interpretation of these multipliers

(given by Bacharach 1970, based on Stone 1962) is to view them as "substitution effects" and "fabrication effects." The row, or substitution effects (r_i 's), measures the

extent that the output of the i-th sector has been substituted for or replaced by other commodities as an input in the production process. The column, or fabrication effects (s_j 's), measures the changes in the relative

weight of value added items in the j-th sector's production. Accordingly, if the output of industry A replaces the output of industry B in the production process, then all the coefficients in the A-th row will go up and all the coefficients in B-th row will decrease. Similarly, if industry C purchases more value added items and less interindustry inputs, all the coefficients in the C-th column will decline. Hence the substitution and fabrication effects can be viewed as reflecting technological changes, changes in relative prices, etc.

4.4) RECTANGULAR RAS METHOD:

In the RAS method, value-added coefficients are treated as residuals, i.e., relying on the basic I-O relationship, the value added vector of the updated year is determined as:

(4-57) $$V_j^t = 1 - \sum_{i=1}^{n} a_{ij}^t$$

for non negative values of a_{ij}^t and V_{ij}^t , where (t) designates the updating year. This residual treatment of the value added vector, however, can be avoided if one incorporates the base year's value added information into the procedure. Then, both a_{ij}^t and V_j^t can be projected via RAS technique. In other words, instead of estimating A^t and then determining V_j^t as the residual, the entire rectangular matrix $\frac{A^t}{V^t}$ will be projected by utilizing the RAS method. The hope is to obtain more reflective estimation of value added items in the updated table.

This procedure, denoted as RECRAS, was first suggested by the Research Institute for Investment of the Japan Development Bank (1977), and to this researcher's knowledge has never been applied anywhere else. The method formally can be presented as:

(4-58) $$\begin{bmatrix} \tilde{A}^t \\ \hline \tilde{V}^t \end{bmatrix} = [\hat{R}] \begin{bmatrix} A^0 \\ \hline V^0 \end{bmatrix} [\hat{S}]$$

Subject to:

(4-59) $$u_i^t = \sum_j^n r_i \, x_{ij}^0 \, s_j$$

and:

(4-60) $$x_j^t = \sum_i^n r_i \, x_{ij}^0 \, s_j$$

where (0) and (t) refer to the base and target year. The matrices $\hat{R}$ and $\hat{S}$ denote the substitution and fabrication

matrices of size (n+1) by (n) and (n) by (n) respectively. The bracket on the left hand side of the equation (4-58) and the middle bracket on the right hand side are (n+1) by (n) rectangular matrices of coefficient and value added for the target year (t) and the base year (0).

A problem, however, arises from this procedure. From the basic I-O relationship it is known that in any I-O table the summation of all the coefficients in any given industry plus the ratio of value added to the total output must add up to one or one hundred percent, i.e., one must have:

(4-61) $$\sum_{i=1}^{n} a_{ij}^{t} + v_{j}^{t} = 1$$

The RECRAS procedure, does not guarantee the satisfaction of this relationship. To rectify the problem, a modification of the procedure has been suggested (Kaneko, 1983). The modification can be summarized as follows:

Assume $\hat{R}$ and $\hat{S}$ in equation (4-58) are determined. Then the task is to find $\hat{R}^*$ and $\hat{S}^*$ in such way that, while satisfying the conditions of the equation (4-61), be as close to $\hat{R}$ and $\hat{S}$ as possible. Formally:

(4-62) $$H = [e]' + [\hat{R}^* - \hat{R}]^2 \; [e] + [e]' \; [\hat{S}^* - \hat{S}]^2 \; [e]$$

subject to:

(4-63) $$[e]'[\hat{R}^*] \left[\frac{A^0}{V^0}\right] [\hat{S}^*] = [e]'$$

where [e] and [e]' represent row and column vectors of sum totals of a_{ij} 's plus v_j 's, viz, all their elements are unity. Let:

(4-64) $$[r^*]' = [e]'[\hat{R}^*]$$

(4-65) $$[r]' = [e]'[\hat{R}]$$

and:

(4-66) $$[q^*]' = [e]'[\hat{S}^*]^{-1}$$

(4-67) $$[q]' = [e]'[\hat{S}]^{-1}$$

Then equations (4-62) and (4-63) can be rewritten as:

(4-68) $$H = [r^* - r]'[r^* - r] + [q^* - q]\ [q^* - q]$$

subject to:

(4-69) $$[r^*]'\left[\frac{A^0}{V^0}\right] = [q^*]'$$

Utilizing the Lagrangian technique and denoting the appropriate Lagrangian vector as $[\lambda]'$, the resultant partial derivatives can be stated as:

(4-70) $$\frac{\partial H}{\partial r^{*\prime}} = 2[r^* - r]' + [\lambda]'\left[\frac{A^0}{V^0}\right]' = [0]'$$

(4-71) $$\frac{\partial H}{\partial q^{*\prime}} = 2[q^* - q]' - [\lambda]' = [0]'$$

(4-72) $$\frac{\partial H}{\partial \lambda'} = [r^*]'\left[\frac{A^0}{V^0}\right] = [q^*]'$$

Inserting $[q^*]'$ from equation (4-72) into (4-71) will yield:

(4-73) $$2\left\{[r^*]'\left[\frac{A^0}{V^0}\right] - [q]'\right\} - [\lambda]' = [0]'$$

or:

(4-74) $$[\lambda]' = 2\left\{[r^*]'\left[\frac{A^0}{V^0}\right] - [q]'\right\}$$

substituting for λ from equation (4-74) into (4-70) results:

(4-75) $$2\{[r^*] - [r]\}' + (2)\left\{[r^*]'\left[\frac{A^0}{V^0}\right] - [q]'\right\}\left[\frac{A^0}{V^0}\right]' = [0]'$$

or:

(4-76) $$[r^*]' - [r]' + [r^*]'\left[\frac{A^0}{V^0}\right]\left[\frac{A^0}{V^0}\right]' - [q]'\left[\frac{A^0}{V^0}\right]' = [0]'$$

which can be rearranged as:

(4-77) $$[r^*]' + [r^*]'\left[\frac{A^0}{V^0}\right]\left[\frac{A^0}{V^0}\right]' = [r]' + [q]'\left[\frac{A^0}{V^0}\right]'$$

or:

(4-78) $$[r^*]'\left\{[I] + \left[\frac{A^0}{V^0}\right]\left[\frac{A^0}{V^0}\right]'\right\} = [r]' + [q]'\left[\frac{A^0}{V^0}\right]'$$

which will give:

(4-79) $$[r^*]' = \left\{[r]' + [q]'\left[\frac{A^0}{V^0}\right]'\right\}\left\{[I] + \left[\frac{A^0}{V^0}\right]\left[\frac{A^0}{V^0}\right]'\right\}^{-1}$$

and:

(4-80) $$[q^*] = [r^*]'\left[\frac{A^0}{V^0}\right]$$

where [I] is an identity matrix of the order (n+1).

Given values of $[r^*]'$ and $[q^*]'$, then, $[\hat{R}^*]$ and $[\hat{S}^*]$ can be constructed as follow:

(4-81) $$[\hat{R}^*] = \begin{bmatrix} r^*_{11} & 0 & 0 & . & . & . & 0 \\ 0 & r^*_{22} & 0 & 0 & . & . & 0 \\ 0 & 0 & r^*_{33} & 0 & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & r^*_{(n+1)(n+1)} \end{bmatrix}$$

and:

(4-82) $$[\hat{S}^*] = \begin{bmatrix} q^*_{11} & 0 & 0 & . & . & . & 0 \\ 0 & q^*_{22} & 0 & 0 & . & . & 0 \\ 0 & 0 & q^*_{33} & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & q^*_{(n)(n)} \end{bmatrix}$$

Having $[\hat{R}^*]$ and $[\hat{S}^*]$ and substituting these values in equation (4-58) for $[\hat{R}]$ and $[\hat{S}]$, the desired matrix for

the target year, viz, $\left[\frac{\tilde{A}^t}{\tilde{V}^t}\right]$, can be obtained. Via this

procedure, then, the value added vector, instead of being treated as residual, is directly updated along with the coefficient matrix.

4.5) MATHEMATICAL PROGRAMMING APPROACH:

It has been suggested that application of the RAS method tends to introduce an upward bias into the projection of the I-O coefficients (Lecomber, 1975). This is due to the fact that in the RAS procedure rows and columns are assumed to be influenced uniformly by the substitution and fabrication effects. The assumption, in turn, leads to exponential projection of each individual cell which could mean an upwardly biased estimation of the coefficients in the target year. On the other hand, researcher in the field of updating I-O matrices have been experimenting with a series of mathematical programming techniques for updating purposes. The Lagrangian multiplier method is one such technique, which potentially can remedy the upward bias of the RAS procedure.

The problem, in this context, is to project a matrix of a future date, say $\tilde{X}$, given a base year matrix of the same dimension, say X , in such a way to satisfy the known marginal column and row totals of the projected matrix $\tilde{X}$. The ideal case is when $\tilde{X}$ is exactly equal to X . In the absence of such a situation, the next best solution is to find an $\tilde{X}$ to be as close as possible to X while conforming to the row and column total constraints. The vectors of these constraints can be expressed as:

(4-83) $$U = Xe$$

and

(4-84) $$V = e'X$$

Since U and V are known in advance, the problem boils down to defining and minimizing a "closeness" function subject to these constraints. The "closeness" is normally defined in terms of distance between each individual cell of $\tilde{X}$ as compared to its counterpart in X .

Variety of minimands are suggested for this purpose. Almon (1968) suggests $\sum_i \sum_j (\tilde{x}_{ij}-x_{ij})^2$. Matuszewski, et.al. (1964) define their minimand as: $\sum_i \sum_j \frac{|\tilde{x}_{ij}-x_{ij}|}{x_{ij}}$. The proposed minimands by Theil (1967) and Friedlander (1961) are expressed as $\sum_i \sum_j \left(x_{ij} \log \frac{x_{ij}}{\tilde{x}_{ij}}\right)$ and $\sum_i \sum_j \frac{(\tilde{x}_{ij}-x_{ij})^2}{x_{ij}}$ respectively. Stephen (1942), Schneider (1965), Omar (1967), Bacharach (1970), Lecomber (1971), Geary (1973), and Henry (1973), among others, suggest minimands for the same purpose. As noted earlier, the question in all these cases remains the same, that is, to minimize the proposed objective function subject to the imposed marginal totals. The process will yield an estimate of the target year's matrix that is as close as possible to the base year matrix without violating the constraints. This presumably removes the upward bias inherent in the RAS procedure. The minimization is achieved via some mathematical optimization technique, e.g., Quadratic Programming, Linear Programming, Lagrangian Multiplier Technique, etc.

The present study utilizes three of the suggested minimands, namely those of Almon (1968) and two variants of Freidlander (1961). This choice stems from several factors. Among them are basic similarity of most of the proposed minimands, ease and efficiency of computation, and widespread acceptability. The Lagrangian optimization technique have been utilized in all three cases.

4.5.1) <u>Friedlander Approach</u>. As noted before, this method, denoted as FRIED, utilizes the Lagrangian optimization technique to obtain a set of estimates for the coefficients of the target year in such a way to minimize the sum of deviations between the base and the estimated coefficients, while satisfying the known marginal totals. Let $\tilde{x}_{ij}^{t}$ and x_{ij}^{0} represent elements of the projected and

base year matrices respectively. If the sum total of relative discrepancy ratios of these elements are denoted by (Q), the problem may be formally stated as:

$$\text{(4-85)} \qquad \textit{Minimize } Q = \sum_{i}^{n}\sum_{j}^{n} \frac{(\tilde{x}_{ij}^{t} - x_{ij}^{0})^{2}}{x_{ij}^{0}}$$

subject to:

$$\text{(4-86)} \qquad \sum_{i}^{n} \tilde{x}_{ij}^{t} = V_{j}^{t}$$

(4-87) $$\sum_{j}^{n} \tilde{x}_{ij}^{t} = U_i^t$$

where $\tilde{x}_{ij}^{t}$ and x_{ij}^{0} are intermediate inputs in the base and target year, and U_i^t and V_j^t are known marginal totals of the projecting year matrix. The auxiliary function will be:

$$\Phi = \sum_{i}^{n}\sum_{j}^{n} \frac{(\tilde{x}_{ij}^{t} - x_{ij}^{0})^2}{x_{ij}^{0}} + \sum_{i}^{n} \lambda_i \left(u_i^t - \sum_{j}^{n} \tilde{x}_{ij}^{t}\right) + \sum_{j}^{n-1} \mu_j \left(v_j^t - \sum_{i}^{n} \tilde{x}_{ij}^{t}\right)$$

(4-88)

where, λ_i and μ_j are the appropriate Lagrangian multipliers.

It should be noted that in the last term of this equation (j) only goes up to (n-1). This is because when (n) constraints with (i), and (n-1) constraints with (j) are satisfied (i.e., 2n-1 equations), then the 2n-th equation will be satisfied by definition. That is to say the value of $u_{i.n}^{t}$ and hence $\tilde{x}_{i.n}^{t}$ is uniquely defined.

Now let:

(4-89) $$w_{ij} = \tilde{x}_{ij}^{t} - x_{ij}^{0}$$

(4-90) $$l_i = u_i^t - \sum_{j}^{n} x_{ij}^{0}$$

(4-91) $$m_j = v_j^t - \sum_{i}^{n} x_{ij}^{0}$$

By virtue of these equations along with equations **(4-86)** and **(4-87)**, one can write:

(4-92) $$l_i = \sum_{j}^{n} \tilde{x}_{ij}^{t} - \sum_{j}^{n} x_{ij}^{0}$$

and:

(4-93) $$m_j = \sum_{i}^{n} \tilde{x}_{ij}^{t} - \sum_{i}^{n} x_{ij}$$

Or alternatively:

(4-94) $$l_i = \sum_{j}^{n} w_{ij}$$

and:

(4-95) $$m_j = \sum_{i}^{n} w_{ij}$$

for i = 1, 2, ..., n and j = 1, 2, ..., n-1.

Then, equation (4-88) may be rewritten as:

$$\Phi = \sum_{i}^{n}\sum_{j}^{n}\frac{w_{ij}^{2}}{x_{ij}} + \sum_{i}^{n}\lambda_i\left(l_i - \sum_{j}^{n} w_{ij}\right) + \sum_{j}^{n-1}\mu_j\left(m_j - \sum_{i}^{n} w_{ij}\right)$$

(4-96)

Taking partial derivatives with respect to variables and setting them equal to zero will yield:

(4-97) $$\frac{\partial\Phi}{\partial w_{ij}} = \frac{2w_{ij}}{x_{ij}} - \lambda_i - \mu_j = 0$$

(4-98) $$\frac{\partial\Phi}{\partial w_{i.n}} = \frac{2w_{i.n}}{x_{i.n}} - \lambda_i = 0$$

(4-99) $$\frac{\partial\Phi}{\partial\lambda_i} = l_i - \sum_{j}^{n} w_{ij} = 0$$

(4-100) $$\frac{\partial\Phi}{\partial\mu_j} = m_j - \sum_{i}^{n} w_{ij} = 0$$

for i = 1, 2, ..., n and j = 1, 2, ..., n-1.

From equations (4-97) and (4-98), one can obtain:

(4-101) $$2w_{ij} = x_{ij}^{0}\ (\lambda_i + \mu_j)$$

and:

(4-102) $$2w_{i.n} = x_{i.n}^{0}\ (\lambda_i)$$

Defining $s_{ij} = \frac{x^0_{ij}}{2}$ and $s_{i.n} = \frac{x^0_{i.n}}{2}$, equations (4-99) and

(4-100) may be restated as:

$$\text{(4-103)} \qquad l_i = \sum_{j}^{n-1} (\lambda_i + \mu_j)(s_{ij}) + \lambda_i s_{i.n}$$

$$\text{(4-104)} \qquad m_j = \sum_{i}^{n} (\lambda_i + \mu_j)(s_{ij})$$

for i = 1, 2, ..., n and j = 1, 2, ..., n-1.

Simultaneous solution of (4-103) and (4-104) will provide the values of λ_i and μ_j , for there are 2n-1 equations and 2n-1 unknowns. In matrix notation the system may be stated as in (4-105) below, where the left hand side bracket is a (2n-1) by (1) column vector, the middle bracket is a (2n-1) by (2n-1) matrix, and the last bracket is a (2n-1) by (1) column vector.

$$\text{(4-105)} \qquad \begin{bmatrix} L \\ M \end{bmatrix} = \begin{bmatrix} A & S & \bar{\lambda} \\ S' & B & \bar{\mu} \end{bmatrix}$$

The system above may be restated in a more detailed version as the expanded matrix (4-106) below.

(4-106)

$$
\left[\begin{array}{c}
l_1 \\ l_2 \\ \vdots \\ l_n \\ \hline m_1 \\ m_1 \\ \vdots \\ m_{n-1}
\end{array}\right]
=
\left[\begin{array}{cccc|cccc}
\sum_j^n s_{1j} & 0 & \cdots & 0 & s_{11} & s_{12} & \cdots & s_{1,n-1} \\
0 & \sum_j^n s_{2j} & \cdots & 0 & s_{21} & s_{22} & \cdots & s_{2,n-1} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & \sum_j^n s_{nj} & s_{n1} & s_{n2} & \cdots & s_{n,n-1} \\
\hline
s_{11} & s_{21} & \cdots & s_{n1} & \sum_i^n s_{i1} & 0 & \cdots & 0 \\
s_{21} & s_{22} & \cdots & s_{n2} & 0 & \sum_i^n s_{i2} & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
s_{1,n-1} & \cdot & \cdots & s_{n,n-1} & 0 & 0 & \cdots & \sum_i^n s_{i,n-1}
\end{array}\right]
\left[\begin{array}{c}
\lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \hline \mu_1 \\ \mu_2 \\ \vdots \\ \mu_{n-1}
\end{array}\right]
$$

The system **(4-106)** may be stated alternatively as:

(4-107)
$$
\left[\begin{array}{l}
L = A\,\lambda + S\,\mu \\
M = S'\,\lambda + B\,\mu
\end{array}\right]
$$

Solving for λ and μ from equation (4-107) will give:

$$\lambda = A^{-1} \, [L - S \, \mu] \tag{4-108}$$

and:

$$\mu = [B - S' \, A^{-1} \, S]^{-1} \, [M - S' \, A^{-1} \, L] \tag{4-109}$$

From the equation (4-89), it is known that:

$$\tilde{x}_{ij}^{t} = x_{ij}^{0} + w_{ij} \tag{4-110}$$

substituting in this equation for w_{ij} from equations (4-101) and (4-102) will yield:

$$\tilde{x}_{ij}^{t} = x_{ij}^{0} + s_{ij} \, (\lambda_i + \mu_j) \tag{4-111}$$

for the first to (n-1)-th columns, and:

$$\tilde{x}_{i.n}^{t} = x_{i.n}^{0} + s_{i.n}(\lambda_i) \tag{4-112}$$

for the n-th column.

Noting that s_{ij} and $s_{i.n}$ were defined as $\frac{1}{2}x_{ij}^{0}$ and $\frac{1}{2}x_{i.n}^{0}$ respectively, one can get:

$$\tilde{x}_{ij}^{t} = x_{ij}^{0} + (\frac{1}{2}) \, (x_{ij})^{0} (\lambda_i + \mu_j) \tag{4-113}$$

and:

(4-114) $$\tilde{x}^t_{i.n} = x^0_{i.n} + (\frac{1}{2})(x^0_{i.n})(\lambda_i)$$

Or:

(4-115) $$\tilde{x}^t_{ij} = x^0_{ij}\ [1 + (\frac{1}{2})(\lambda_i + \mu_j)]$$

and:

(4-116) $$\tilde{x}^t_{i.n} = x^0_{i.n}\ [1 + (\frac{1}{2})(\lambda_i)]$$

for i = 1, 2, ..., n and j = 1, 2, ..., n-1.

Or in matrix notation:

(4-117) $$[\tilde{X}^t] = [\hat{\lambda}][S] + [S]\left[\frac{\hat{\mu}}{0}\right] + [X^0]$$

where $\tilde{X}^t$ and X^0 are (n) by (n) estimated and actual matrices of the target year and base year respectively, S is an (n) by (n) matrix as defined before, and $\hat{\lambda}$ and $\hat{\mu}$ are (n) by (n) diagonal matrices constructed from column and row-wise adjusting factors: λ_i and μ_j . It should be noted that the value of μ_n is equal to zero, i.e., $x^0_{i.n}$ is subjected only to a row-wise adjustment.

An important observation must be made here with regards to any Lagrangian type method. In some I-O tables there might be some null rows or columns, e.g., an industry that receives from other sectors but its entire output goes to final demand, hence containing zeroes in its row in the inter-industry transaction matrix (in fact the construction industry in the Soviet table for 1972 had a null row). Aside from the fact that such a situation is normally an indication of inappropriate classification, operationally it makes the usage of Lagrangian multipliers impossible. For a null row or column in an I-O table, will render the coefficient matrix singular. Concretely speaking, this means that matrix [A] of the equations **(4-106)** and **(4-107)** or the **(4-105)**, i.e,. the system of equations for obtaining the Lagrangian multipliers λ and μ , will be singular. This singularity, in turn, will make the inversion of this matrix impossible, thereby making the solution to the Lagrangian multipliers non-existent.

This is not a surprising outcome, since a row or column of zeroes implies that the system is overdetermined and a unique solution does not exist. To remedy the situation, one must either start with a better classification at the construction phase of the I-O table, or eliminate the zero vector by an appropriate aggregation. This in essence translates into elimination of enough of the equations, to the point that the system will no longer be overdetermined.

In any event, the elements of the target year's matrix are constructed via utilization of the system of equations (4-115) and (4-116), or (4-117). Each element of the target year matrix is derived through summation of its corresponding element in the base year's matrix and the weighted average of the column-wise and row-wise adjusting factors λ_i and μ_j . The weight attached to these factors is the value of the relevant base year's element.

As mentioned earlier, in the RAS procedure the base year's elements are adjusted by substitution and fabrication multipliers. The process as a whole, and particularly when it concerns the cells with relatively small values, tends to introduce an upward bias into the estimates. The Freidlander approach may provide an alternative to researchers that does not suffer from this defect, thereby yielding a more accurate estimation of the target year's coefficients.

4.5.2) Rectangular Friedlander Approach. The original Friedlander approach, just like the RAS procedure, estimates the interindustry coefficients and treats the value added of each sector as residuals. However, parallel to the case of RECRAS method, one can argue for inclusion of the value added data in the process and obtain these values directly, instead of treating them as residuals. This idea is used by Koneko (1983) and has been utilized in the present study via rectangularization of the Friedlander objective function. The method, designated here as RECLAG, is presented below:

Let R^0 and $\tilde{R}^t$ represent a rectangular (n+k) by (n) matrices of the base and target year respectively. In these matrices, the first (n) elements refer to intermediate inputs and (n+1) to (n+k) elements denote the input of primary factors. In other words, there are (n) industries and (k) primary factors.

Through steps similar to the Friedlander approach, it can be shown that the values of Lagrangian multipliers can be obtained from simultaneous solution of the appropriate partial derivatives of the following function:

(4-118) $$\Phi = \sum_i^{n+k} \sum_j^{n} \frac{w_{ij}^2}{R_{ij}^0} + \sum_i^{n+k} \lambda_i \left(l_i - \sum_j^{n} w_{ij} \right) + \sum_j^{n-1} \mu_j \left(m_j - \sum_i^{n+k} w_{ij} \right)$$

where $l_i = \sum_j^{n} w_{ij}$, $m_j = \sum_i^{n+k} w_{ij}$, $\tilde{R}_{ij}^t - R_{ij}^0 = w_{ij}$, and $\tilde{R}_{ij}^t$ and R_{ij}^0 are the elements of the target year and the base year matrices. The Lagrangian multipliers, as before, are denoted by λ_i and μ_j .

The resultant system here, contains (2n+k-1) equations and the same number of unknowns, hence it can be solved. The (2n+k)-th equation is, then, uniquely defined as well.

Defining $S_{ij} = (\frac{1}{2})(R^0_{ij})$, the system can be presented in

matrix notation as **(4-119)** below:

(4-119)

$$
\begin{bmatrix} l_1 \\ l_2 \\ \cdot \\ \cdot \\ \cdot \\ l_{n+k} \\ \hline m_1 \\ m_2 \\ \cdot \\ \cdot \\ m_{n-1} \end{bmatrix}
=
\left[\begin{array}{ccccc|ccccc}
\sum_j^n S_{1j} & 0 & \cdot & \cdot & 0 & S_{11} & S_{12} & \cdot & \cdot & S_{1,n-1} \\
0 & \sum_j^n S_{2j} & \cdot & \cdot & 0 & S_{21} & S_{22} & \cdot & \cdot & S_{2,n-1} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdot & \cdot & \sum_j^n S_{n+k,j} & S_{n+k,1} & S_{n+k,2} & \cdot & \cdot & S_{n+k,n-1} \\
\hline
S_{11} & S_{21} & \cdot & \cdot & S_{n+k,1} & \sum_i^{n+k} S_{i1} & 0 & 0 & \cdot & 0 \\
S_{12} & S_{22} & \cdot & \cdot & S_{n+k,2} & 0 & \sum_i^{n+k} S_{i2} & \cdot & \cdot & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
S_{1,n-1} & \cdot & \cdot & \cdot & S_{n+k,n-1} & 0 & 0 & \cdot & \cdot & \sum_i^{n+k} S_{i,n-1}
\end{array}\right]
\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_{n+k} \\ \hline \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \mu_{n-1} \end{bmatrix}
$$

Having the value of the Lagrangian multipliers, as have been shown previously, the elements of the target year matrix can be found for j=1,2,...,n-1 and i=1,2,...,n+k as:

$$\tilde{R}_{ij}^{t} = R_{ij}^{0}\left[1 + \left(\frac{1}{2}\right)(\lambda_i + \mu_j)\right] \tag{4-120}$$

and for the n-th vector as:

$$\tilde{R}_{i.n}^{t} = R_{i.n}^{0}\left[1 + \left(\frac{1}{2}\right)(\lambda_i)\right] \tag{4-121}$$

In other words, the same solution procedure and format as the regular Friedlander approach is used, except for the fact that information on value added is included and estimation has been conducted accordingly.

4.5.3) **Almon Approach**. This method is another Lagrangian type method, hence utilizes the Lagrangian optimization technique. The logic and explanation is the same as before. The difference, however, is in the proposed minimand. Almon (1968) suggests the objective function as:

$$\textit{Minimize} \quad \sum_{i}^{n}\sum_{j}^{n}(\tilde{x}_{ij}^{t} - x_{ij}^{0})^{2} \tag{4-122}$$

subject to:

$$\sum_{j}^{n}\tilde{x}_{ij}^{t} = u_{i}^{t} \tag{4-123}$$

and:

(4-124) $$\sum_{i}^{n} \tilde{x}_{ij}^{t} = v_j^t$$

where $\tilde{x}_{ij}^{t}$ and x_{ij}^{0} are the elements of the projecting and base year matrix respectively, and u_i^t and v_j^t denote the known marginal totals of the updating year's matrix.

Utilizing the Lagrangian technique, the auxiliary function will be:

(4-125)

$$\Phi = \sum_{i}^{n}\sum_{j}^{n} (\tilde{x}_{ij}^{t} - x_{ij}^{0})^2 + \sum_{i}^{n} \lambda_i (u_i^t - \sum_{j}^{n} \tilde{x}_{ij}^{t}) + \sum_{j}^{n-1} \mu_j (v_j^t - \sum_{i}^{n} \tilde{x}_{ij}^{t})$$

where λ_i and μ_j are the Lagrangian multipliers and other terms are the same as before.

It should be noted that, here too, only solution of (2n-1) equations suffice to obtain (2n) uniquely defined answers to the unknowns.

Now let:

(4-126) $$\tilde{x}_{ij}^{t} - x_{ij}^{0} = w_{ij}$$

(4-127) $$u_i^t - \sum_j^n x_{ij}^0 = l_i$$

(4-128) $$v_j^t - \sum_i^n x_{ij}^0 = m_j$$

Then:

(4-129) $$l_i = \sum_j^n w_{ij}$$

(4-130) $$m_j = \sum_i^n w_{ij}$$

for i = 1, 2, ...,n and j = 1, 2, ..., n-1.

Thus, the auxiliary function can be rewritten as(4-131) below:

(4-131)

$$\Phi = \sum_i^n \sum_j^n (w_{ij})^2 + \sum_i^n \lambda_i \left(l_i - \sum_j^n w_{ij}\right) + \sum_j^{n-1} \mu_j \left(m_j - \sum_i^n w_{ij}\right)$$

Taking partial derivatives of this function with respect to its variables and setting them equal to zero will yield:

(4-132) $$\frac{\partial \Phi}{\partial w_{ij}} = 2w_{ij} - \lambda_i - \mu_j = 0$$

for j = 1, 2, ..., n-1, and:

(4-133) $$\frac{\partial\Phi}{\partial w_{i.n}} = 2w_{i.n} - \lambda_i = 0$$

for j = n, as well as:

(4-134) $$\frac{\partial\Phi}{\partial\lambda_i} = l_i - \sum_j^n w_{ij} = 0$$

(4-135) $$\frac{\partial\Phi}{\partial\mu_j} = m_j - \sum_i^n w_{ij} = 0$$

for the Lagrangian multipliers. These partial derivatives may be written as a system, i.e.,:

(4-136) $$\begin{aligned} 2w_{ij} &= \lambda_i + \mu_j \\ 2w_{i.n} &= \lambda_i \\ l_i &= \sum_j^n w_{ij} \\ m_j &= \sum_i^n w_{ij} \end{aligned}$$

Which yields:

(4-137) $$\begin{aligned} l_i &= (\tfrac{1}{2})\sum_j^{n-1} (\lambda_i + \mu_j) + (\tfrac{1}{2})\lambda_i \\ m_j &= (\tfrac{1}{2})\sum_i^{n} (\lambda_i + \mu_j) \end{aligned}$$

or:

(4-138) $$\begin{aligned} 2l_i &= \sum_j^{n-1} (\lambda_i + \mu_j) + \lambda_i \\ 2m_j &= \sum_i^{n} (\lambda_i + \mu_j) \end{aligned}$$

for i = 1, 2, ..., n and j = 1, 2, ..., n-1.

In matrix notation, The system can be shown as (4-139) below:

(4-139)

$$\begin{bmatrix} 2l_1 \\ 2l_2 \\ \cdot \\ \cdot \\ \cdot \\ 2l_n \\ \hline 2m_1 \\ 2m_2 \\ \cdot \\ \cdot \\ 2m_{n-1} \end{bmatrix} = \left[\begin{array}{cccccc|cccccc} n & 0 & . & . & 0 & & 1 & 1 & . & . & 1 & \\ 0 & n & . & . & 0 & & 1 & 1 & . & . & 1 & \\ . & . & . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . & . & . \\ 0 & . & . & . & n & & 1 & 1 & . & . & 1 & \\ \hline 1 & 1 & . & . & 1 & & n & 0 & . & . & 0 & \\ 1 & 1 & . & . & 1 & & 0 & n & . & . & 0 & \\ . & . & . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . & . & . \\ 1 & 1 & . & . & 1 & & 0 & 0 & . & . & n & \end{array}\right] \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_n \\ \hline \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \mu_{n-1} \end{bmatrix}$$

Or, alternatively:

(4-140)
$$\begin{aligned} L &= A\,\lambda + W\,\mu \\ M &= W'\,\lambda + B\mu \end{aligned}$$

Simultaneous solution of these equations will provide the values for the Lagrangian multipliers λ and μ , i.e.,:

(4-141)
$$\begin{aligned} \lambda &= A^{-1}\,[\,L - W\,\mu\,] \\ M &= W'\,A^{-1}\,[\,L - W\,\mu\,] + B\,\mu \end{aligned}$$

Substituting for λ from the first set into the second will give:

(4-142) $$\begin{aligned} \overline{\lambda} &= A^{-1} [L - W \overline{\mu}] \\ M &= W' A^{-1} L - W' A^{-1} W \overline{\mu} + B \overline{\mu} \end{aligned}$$

or:

(4-143) $$\begin{aligned} \lambda &= A^{-1} [L - W \mu] \\ B \mu - W' A^{-1} W \mu &= M - W' A^{-1} L \end{aligned}$$

Which may be rearranged as:

(4-144) $$\begin{aligned} \lambda &= A^{-1} [L - W \mu] \\ [B - W' A^{-1} W] \mu &= M - W' A^{-1} L \end{aligned}$$

and leads to:

(4-145) $$\begin{aligned} \overline{\lambda} &= A^{-1} [L - W \overline{\mu}] \\ \overline{\mu} &= [B - W' A^{-1} W]^{-1} [M - W' A^{-1} L] \end{aligned}$$

Having λ and μ , w_{ij} 's and thus $\tilde{x}^{t}_{ij}$'s are determined

since x^{0}_{ij} 's are already known. Formally:

(4-146) $$2 w_{ij} = \lambda_i + \mu_j$$

or:

(4-147) $$w_{ij} = (\frac{1}{2}) (\lambda_i + \mu_j)$$

since:

(4-148) $$\tilde{x}^{t}_{ij} = x^{0}_{ij} + w_{ij}$$

therefore:

(4-149) $$\tilde{x}_{ij}^{t} = x_{ij}^{0} + \frac{\lambda_i + \mu_j}{2}$$

for the first through (n-1)-th elation , and:

(4-150) $$\tilde{x}_{i.n}^{t} = x_{i.n}^{0} + \frac{\lambda_i}{2}$$

for the n-th equation. Or in matrix notation:

(4-151) $$[\ \tilde{X}^{t}\ ] = (\frac{1}{2})\{\ [\ \hat{\lambda}\ ] + [\ \mu\ ]\ \} + [\ X^{0}\ ]$$

The solution implies that the Almon approach amounts to adjustment of each cell in the base year matrix by a simple average of its respective substitution and fabrication effects, in order to arrive at the corresponding cell for the target year matrix.

4.6) COMBINED RAS-LAGRANGIAN APPROACH:

As noted earlier, usage of RAS and RECRAS techniques tend to introduce an upward bias into the coefficients of the target year matrix, particularly in the case of coefficients with small values in the base year. Utilization of Lagrangian technique in place of RAS or RECRAS method was one attempt to rectify this problem. Another method utilized by Koneko (1983), is combination of the two methods or what he calls Two stage RAS-Lagrangian method, and is denoted here as RASLAG.

Under this method, the target year matrix of coefficients is first estimated via RAS method. Then the Lagrangian optimization technique is applied to RAS generated results in order to get the final adjusted estimates for the target year. The minimand used here, is that of Friedlander. In other words at first, the coefficients are obtained through:

(4-152) $$\tilde{A}^t = \hat{R}\, A^0\, \hat{S}$$

subject to:

(4-153) $$u_i^t = \sum_j^n r_i\, x_{ij}^0\, s_j$$

and:

(4-154) $$v_j^t = \sum_i^n r_i\, x_{ij}^0\, s_j$$

for i and j = 1, 2, ..., n.

Where x_{ij}^0 is an element of the base year's interindustry transaction matrix , u_i^t represents the predetermined value of intermediate demand for the i-th commodity, and v_j^t denotes the known intermediate inputs of j-th industry, in the target year. $\tilde{A}^t$, A^0 , $\hat{R}$,and $\hat{S}$ refer to the target year matrix, base year matrix, and diagonal matrices of substitution and fabrication effects respectively.

Having the RAS generated estimates of the projecting year's matrix of coefficients, (i.e., $\tilde{A}^t$), the matrix of intermediate inputs can be constructed by:

$$\text{(4-155)} \qquad \tilde{X}^t = \tilde{A}^t \hat{X}_j^t$$

where $\hat{X}_j$ is the diagonal matrix of total output of the projecting year (t). Now, in a process similar to what has been explained previously, the Lagrangian technique can be applied to this RAS generated estimates in order to rectify any possible upward bias in the coefficients. In other words, the RAS generated, and Lagrangian adjusted estimate of the matrix of coefficients for target year, (i.e., $\overline{A}^t$), is obtained via:

$$\text{(4-156)} \qquad \overline{x}_{ij}^t = (\tilde{x}_{ij}^t) \left\{ (1) + (\tfrac{1}{2}) (\lambda_i + \mu_j) \right\}$$

Or in matrix notation:

(4-157) $\bar{X}_t = [\hat{\lambda}]\,[S] + [S]\,[\hat{\mu}] + [\tilde{X}^t]$

where S is defined as $(\frac{1}{2})\tilde{x}^t$, and $\hat{\lambda}$ and $\hat{\mu}$ are diagonal

matrices of Lagrangian adjusting multipliers.

Obtaining the coefficients from the resultant matrix of intermediate transactions is a simple matter of ratio calculation. This two-step process, using the same amount of information, potentially may produce more accurate results than the application of RAS procedure alone.

4.7) COMBINED RECTANGULAR RAS-LAGRANGIAN APPROACH:

This is the application of the two stage process to the results generated by RECRAS approach instead of RAS method, and denoted in the current study as RERALA. Under this method, first the base year's matrix and the target year's marginal totals will be used via RECRAS method to generate the matrix of coefficients for the target year, viz:

(4-158) $$\left[\frac{\tilde{A}^t}{\tilde{V}^t}\right] = [\hat{R}]\left[\frac{A^0}{V^0}\right][\hat{S}]$$

subject to:

(4-159) $$u_i^t = \sum_j^n r_i \, x_{ij}^0 \, s_j$$

and:

(4-160) $$X_j^t = \sum_i^n r_i \, x_{ij}^0 \, s_j$$

where all notations are as before.

Having the RECRAS generated matrix of coefficients for the target year, then one can construct the matrix of intermediate inputs and value added via:

(4-161) $$\left[\frac{\tilde{X}^t}{\tilde{V}^t}\right] = \left[\frac{\tilde{A}^t}{\tilde{V}^0}\right] \left[\hat{X}_j^t\right]$$

The Lagrangian technique, then can be applied to this transaction matrix, to obtain the RECRAS generated, and Lagrangian adjusted interindustry matrix. Each element of this matrix, as shown before, can be found through:

(4-162) $$\left(\frac{\overline{x}_{ij}}{\overline{\overline{v}}_j}\right) = \left\{ 1 + \left(\frac{1}{2}\right) (\lambda_i + \mu_j) \right\} \left(\frac{\tilde{x}_{ij}^t}{\tilde{v}_j^t}\right)$$

From the resultant interindustry matrix, then, the matrix of coefficients for the target year is obtained. This procedure also may eliminate the upward bias inherent in the RECRAS method, hence providing a more accurate estimate of the coefficients. It is evident that the minimand utilized here is the one suggested by Friedlander.

4.8) PROPORTIONAL TO VALUE ADDED APPROACH:

This method revolves around a basic assumption with regards to the structure of an I-O table. It assumes that all transactions are proportional to value added, i.e., in the target year, the amount of industry (i)'s output that is used by industry (j) as its input, is proportional to the value added by the industry (j). This assumption implies that the usage of any input into the production process of each commodity has a fix proportion to the amount of labor and capital used in that process. Formally:

$$\psi_{ij} = \frac{x_{ij}^0}{v_j^0} \tag{4-163}$$

where x_{ij}^0 is the amount of industry (i)'s output that is used by industry (j) as input, and v_j^0 refers to the value added of the j-th sector. Denoting the matrix of these proportions by Ψ , the target year's interindustry demand can be obtained via:

$$\tilde{X}^t = [\Psi^0][\hat{v}_j^0] \tag{4-164}$$

Adding the final demand vector to these intermediate demands will yield the gross output of the target year. This technique is included in the present study for its simplicity and plausibility, and denoted here as PROPVA.

4.9) **NAIVE METHOD**:

This method, designated as NAIVE, assumes no change in the I-O coefficients between the base year and the target year tables. In other words it adopts the basic I-O assumption of intertemporal stability of coefficients by simply applying the base year coefficients to the target year. It is included in this study for its simplicity, as well as a mean to test the intertemporal stability assumption via comparison of the NAIVE results with other updating techniques.

The afformentioned methods will be applied to the data and accordingly nine estimated matrices will be obtained. In the next step, these results will be compared with the benchmark table as well as with one another in order to evaluate and rank the updating methods. Next two chapters are devoted to this task.

CHAPTER FIVE

IMPLEMENTATION

"An unsophisticated forecaster uses statistic
as a drunken man uses lamp-posts
-for support rather than for illumination."

Andrew Lang

Prior to engaging in the simulation procedures, some preliminary tasks must be accomplished. The data should be explained and justified, the goals of the project as well as the simulation procedure must be specified, and methodology of obtaining and evaluation of the results ought to be defined. This chapter is designed to address these and related questions.

5.1) DATA:

For experimental purposes, the present study utilizes the Soviet Union's Input-Output tables of 1966 and 1972, expressed in current producers' price. The Soviet authorities, as a matter of policy, did not publish completed I-O tables at one time. Instead, separate blocks of these tables were released at different times and on various occasions. These released blocks, along with variety of other available information, then, were put together in order to arrive at the complete tables.

Currently, there are three such tables available in the West. Namely, 1956, 1966, and 1972 tables. The tables used here are those compiled at the Foreign Demographic Analysis Division of the Bureau of the Census of the U.S. Department of Commerce by Kostinsky, et al. (1976) and Gallik, et al. (1983). The original 1966 Soviet table is an ex-post one, expressed in terms of current purchasers' price, and consists of 131 rows and columns and three quadrants. The interindustry sector represents 110 productive sectors, and there are 21 consuming sectors and 21 value added items. These data are accompanied by data for employment and capital stock, and all tables of direct and total coefficients are derived from them. The original 1972 table is also an ex-post table expressed in purchasers' price. and consists of three quadrant and 112 productive sectors. The exact numbers of final demand and value added quadrants are not known. Based on the previous tables and some additional information, the number of these sectors are estimated to be 15-20 columns of final demand and 7-10 rows of value added plus a vector of depreciation (Gallik, et al. 1983).

Due to lack of reliable data, however, these original tables had to be aggregated in the reconstruction process. These tables are provided in two value variants, namely in terms of Producers' price and Purchasers' price. The resultant tables have the following characteristics. The 1966 table contains 75 producing sectors, five columns of final demand, and three rows of value added, along with a row vector of depreciation and two vectors for fixed capital stock and employment. The 1972 table consists of 88 producing sectors, six final demand columns, six value added rows, along with a row for depreciation and supplemental rows of fixed capital stock and employment.

For the purpose of current project, however, these tables had to be further aggregated to make them operationally compatible. The final tables contain 71 producing sectors, one column vector of final demand, and one row vector of value added. No use of the employment and fixed capital data has been made. Description of the sectoral classification of these tables and their corresponding sectors in the "reconstructed" and "original" tables, along with the aggregation scheme used in this project, are given in Appendix (G).

At this point, it seems necessary to briefly explain the logic behind selection of the Soviet tables in general as well as choosing the 1966 and 1972 current producers' price variants in particular. The issues of reliability of the tables and specific dates also must be addressed.

In order to carry out this study, a country or a region with actual input-output tables must be selected. This requirement immediately eliminates many countries and regions. The choice is further limited due to the fact that at least two such tables must be present for comparative purposes of the current research. From the remaining list, regional tables are eliminated, primarily for three reasons. First, the focus of the present researcher is on the national, rather than regional, economic policies. Second, in most cases all of the available regional tables are not survey based. That is, one survey based table exists and subsequent tables are updated versions of the same table (e.g., Kansas, scotland, Philadelphia, Nebraska, etc.). Third, in the cases where more than one actual table exists, e.g., West Virginia, Washington, Italy-regional, Japan-regional, Netherlands-regional, etc., the tables are either highly aggregated or "overused". In "highly aggregated tables," the level of aggregation is so high that renders them practically useless for the purpose of this study. This is because as aggregation level increases, the tables tend to reflect overall instead of individual tendencies, hence making a cell by cell comparison of the tables less meaningful. The other category, the "overused," is dropped as a candidate for empirical work because there is already a large volume of research conducted on these tables, covering practically every conceivable aspect.

Same problems exist for many national tables. Few relatively abundant national tables that are somewhat reliable exist, e.g., Japan, Netherlands, USA, etc. These tables are generally "overused." There are number of countries that have I-O tables e.g., Iran, Brazil, Pakistan, Korea, etc. These tables, however, are either updated versions of survey based tables or are not very reliable.

It is consideration of these points that led to selection of the Soviet tables. There are three survey based actual I-O tables available, which make the experiment enticing. Not much work is done on the Soviet data, which makes the investigation attractive. The tables are product of centrally planned economic process, which makes the conditions almost laboratory like and desirable.

The question of reliability of the Soviet tables is not entirely satisfied. However, these tables as noted by Kostinsky, et al., (1976) and Gallik, et al., (1983), are compiled from variety of sources and for the most part have been through cross checks. As such, they are the most complete and reliable Soviet I-O tables available to date, and readily lend themselves to appropriate analytical studies. Furthermore, in the context of present research, the data is even more reliable. For what is examined here is the efficiency of various techniques in the generation of non-survey I-O tables, and no structural or economic conclusion about the Soviet Union itself have been made. In other words, so long as the two matrices used here are consistent, even though they may contain "false" or "distorted" data, the comparison of estimation techniques can be safely carried out and the results can be viewed as reliable as the case in which there were no distortion or falsification of data.

From of the three available tables, i.e., 1956, 1966, and 1972, the latter two are chosen for this project. The tables selected here expressed in current, as opposed to constant prices, and producers', rather than purchasers' prices. There are several reasons for these choices, which are explained below.

The change over time of the coefficients of I-O tables are virtually accepted universally. Thus, the results obtained via utilization of the latest survey based tables are less than reliable, particularly as the time gap between the date of the table and the date of analysis widens. Therefore updating the tables through some method is advisable. Finding an efficient method for such purpose, is one of the main objectives of the current project. If a reasonably accurate results could be obtained through some method, then I-O analysts will have reasonable basis to conduct researches between publication of two actual survey based tables. Since the construction of a survey based table takes anywhere between five to ten years, an updating method that can provide reasonable results in this time frame is the most desirable one. Of course, longer time frames may be used for various theoretical inquiries, but considering that the objective here is mainly a practical one, five to ten years time span appears to be optimal. The implicit assumption, obviously is that actual tables are continuously being constructed. Hence, the best choices for this project are pairs of either 1956 and 1966 or 1966 and 1972. The first pair, however, encompasses some period that is marked with major changes and restructuring in the Soviet economy associated with the Kosygin-Liberman reform of 1963. Selection of this time frame will introduce some elements of change in the tables that in actuality have nothing to do with changes in I-O coefficients or updating techniques, and unnecessarily disturb the situation. 1966- 1972 period, on the other hand, is associated with relative stability. Thusly, the 1966 and 1972 tables are adopted for the experimental purposes.

As to the selection of producers' price version over the purchasers' price variant, suffice to mention that the retail prices in the Soviet Union for the most part are policy tools and not reflection of the market forces. As such, changes in the coefficients calculated based on purchasers' price may not be due to substitution of inputs, changes in relative prices of inputs, product mix changes, etc. Producers' price table is far less policy driven and more representative of the true economic conditions governing the inputs and outputs in the Soviet economy.

The last point to be addressed is selection of current price over constant price table. Leontief (1951) expressed his model in terms of physical quantities. The interindustry coefficients, equation (2-7) above, were expressed as ratios of the physical quantity of output of one sector, (i), used as an input in another industry (j). The production theory behind these coefficients implies linear production functions with fixed proportions among the inputs. These physical coefficients are assumed to be fixed for each sector. In practice, however, the data is collected and tables are expressed in monetary term. Thus, to express tables in quantity term in any year other than the base year, one should deflate the monetary expression, i.e., tables must be expressed in constant prices. Conversion of tables expressed in current prices to constant prices, however, requires utilization of price deflators. Irrelevance of Soviet Union's price deflators notwithstanding, the procedure requires either availability of sectoral price indices or utilization of general price index. The first solution is not feasible due to lack of data availability, and the second route introduces unnecessary errors into the coefficients.

Furthermore, fixed coefficients assumption indicates each sector produces a single homogeneous output. But as noted by researchers such as Klein (1953) and Bezdek (1984), joint products are produced in many industries. Therefore, the numerator and denominator of equation (2-7) are not quantity of commodity (i) used as input in production of commodity (j) over the total production of commodity (j). Instead they are weighted averages of several inputs classified in sector (i) that are used as input for production of several commodities classified in industry (j). The weights of each commodity, then, is determined by its price and quantity. Bezdek (1984) argues that Leontief's assumption implies that coefficients remain constant so long as the tables are expressed in constant prices. Klein (1953), suggests that this requires zero price elasticity of substitution between inputs. On the other hand, viewing the I-O coefficients as weighted average of several commodities indicates that for coefficients to stay constant the changes in relative prices must generate offsetting changes in quantities. This indicates unitary price elasticity of substitution among the inputs. Klein's proposition is taken by I-O analysts such as Bezdek and Wending (1976) and Morishima (1956) to mean greater intertemporal stability for tables expressed in current prices. This proposition have been supported by empirical studies conducted, among others, by Tilanus (1966), Czamanski and Maizia (1969), Vaccara (1969), Carter (1970), and Bezdek (1984).

Theoretical discussion and the empirical results cited above are the main rationale behind the selection of the current price version of the Soviet Input-Output tables in this study. Additionally, the results of researches mentioned above, indicate that as the level of aggregation increases, the degree of stability of the coefficients expressed in constant prices becomes even greater. Since experiments with several aggregated tables are included in the present project, it seems logical to work with inherently more stable tables, i.e., tables expressed in current prices. Thus, in light of the above mentioned factors, the most logical and potentially useful approach is taken, namely employment of current price variant of the Soviet Input-Output tables.

5.2) EXPERIMENTAL PROCEDURE:

In this study, adjusted and operationally compatible survey-based I-O tables of the Soviet Union for 1966 and 1972 are used for the simulation purposes. Each table is expressed in producers' price with a 71 by 71 interindustry matrix. The 1966 table is used as the base table and 1972 is chosen as the benchmark table that measures the accuracy of the estimates.

The goal is to find the most efficient projection of 1972 table through a non-survey technique. Efficiency in this context refers to the trade-off between time and effort (cost) and accuracy. Projections of direct and inverse transaction matrices of 1972 table will be made via utilization of eight different updating techniques. These acronymed techniques are: RAS, RECRAS, PROPVA, FRIED, RECLAG, RASLAG, RERALA, and ALMON. Details of these methods were given in chapter four and will not be repeated here. The eight projections along with NAIVE or the 1966 coefficients (assumption of no change in the coefficients between 1966 and 1972) will be compared with the benchmark table, i.e., the actual survey based 1972 coefficients. The comparison should reveal the ranking of the updating techniques and evaluate the justification for their usage.

The next step is to make use of available additional information. The techniques used in the first step are minimum information methods. However, usually additional data about the target year are available and could be used. Moreover, due to host of reasons, researchers might be interested to separately estimate a set of the target year coefficients. Under these circumstances, the updating techniques must be modified to accommodate the inclusion of additional exogenous information.

To evaluate the modification methods, as well as their feasibility and justification, ten different estimates of direct and inverse coefficients of the 1972 table are obtained and compared. The methods used to procure these estimates are: RESMIN, NAIVE, Lagrangian, and RAS. These schemes are chosen for their logical plausibility, wide spread acceptability, relative superiority of performance, and operational ease. A detailed explanation of these modified procedures can be found in chapter seven and will not be pursued here.

The residual Minimum Method (RESMIN) does not incorporate in itself any additional data about the target year matrix. Instead, it estimates the coefficients through Lagrangian minimization procedures while assigning a weight to each coefficient. The other nine estimates are three variants of NAIVE, RAS, and Lagrangian methods.

The NAIVE method, or constant coefficient assumption, is selected to verify whether updating techniques perform better than simply incorporating some additional data into the most recent survey based table or not. The results also can be used as some indication of temporal stability of the I-O coefficients.

From the survey of literature as well as the results obtained in this study, it is evident that application of the simple RAS procedure produces estimates that are superior to other techniques. Also, it has been suggested that in most cases utilization of additional data improves the RAS generated estimates. To verify this hypothesis and due to its superior performance, the RAS method was included in the modification scheme. The Lagrangian method (under Friedlander minimand) performed well in the first round of experiments conducted in this study, and it is also very logical and conceptually sound. Therefore it was also selected for modification purposes.

Each of the above methods was modified under three different scenarios. The selection criterions for incorporation of additional information were "key sectors," "large coefficients," and "most important parameters." The logic, justification, and explanation of the selection criterions are deferred to the next chapter.

The first scenario assumes knowledge of some key sectors. The key sectors chosen in this project were energy sectors. The logic was that a) most countries have separate data on energy, b) these data are much more up to date than I-O tables, c) these data are normally reliable, and d) usually the energy sectors are either state monopolies or closely watched private oligopolies. In either case, it is theoretically much easier to obtain data and forecasts about them. It must be noted that this choice is judgmental and under different conditions a different set of sectors can be chosen with an equally plausible argument. The choice of sectors, however, does not change the procedures.

In the Soviet tables, seven rows (rows 5-11, containing 497 coefficients) constitute the energy sector. It was assumed that all these coefficients for 1972 were known along with the marginal totals of that year's table. In practice, these values must be determined exogenously and in advance. But for the purpose of this research, since the actual 1972 table was available and the aim was to asses relative accuracy of different updating techniques, the exact values, instead of estimates, were taken.

The Second scenario was to choose a given percentage of the entire coefficients, based on some criterion, and subject them to exogenous estimation. For the sake of continuity and ease of comparison, 497 coefficients were selected (equal to number of coefficients of energy sectors selected in the first scenario), which is approximately 10% (9.8592%) of the total set of coefficients in the 1972 table. These were the largest 497 coefficients of that year. The reason for choosing the largest coefficients criterion was twofold. First, as it is indicated by several researchers in the field (see chapter seven) the largest coefficients exert the most influence on the I-O table, hence their accuracy is more crucial to obtain a better estimate. Second, it is usually easier to obtain reliable data on large interindustry transactions than the small ones. The actual values of these 497 largest coefficients in the 1972 table is taken to be their exogenous estimates.

Again, it goes without saying that this choice does not affect the comparison of various techniques.

The third and last scenario is based on selection of "the most important" parameters in the model, i.e., those coefficients whose change (or inaccuracy) will cause the largest number of other coefficients to be disturbed and hence their impact is most extensive and damaging to the system as a whole. Here too, for sake of continuity and comparison, the 497 most important parameters were chosen for exogenous estimation. Similar to the previous scenarios, the actual 1972 values were used instead of their estimates.

Three sets of (497 each) coefficients chosen via the above selection criterions were used to modify the I-O tables and the resultant estimates were recorded. First, it was assumed that exogenously estimated values of 497 key coefficients (energy sectors in this case), 497 largest coefficients, and 497 most important coefficients were obtained. The other coefficients in the table were assumed constant (i.e., the same as 1966). The three generated estimates are labelled as MDNAVKEY, MDNAVBIG, MDNAVMIP respectively. Then, the same three sets of independently estimated coefficients were introduced to the Lagrangian and RAS methods, with one difference. That is, unlike in the NAIVE case, the remaining coefficients were not assumed constant. Instead they were estimated via utilization of Lagrangian and RAS techniques. The combination of these

Lagrangian (or RAS) generated coefficients along with the three sets of predetermined coefficients yielded six additional complete estimated table for 1972. These estimates are designated as MDLAGKEY, MDLAGBIG, MDLAGMIP, MDRASKEY, MDRASBIG, and MDRASMIP respectively.

Before closing this section, some explanation and clarification of few points, are in order. First, as mentioned before, if the [A] matrix in the solution to Lagrangian multipliers (e.g., equations **4-108** and **4-109** of chapter four) or its associated partitioned matrix (e.g., the square matrix of the equation **4-105** in chapter four) contain any null row or column, its determinant will equal zero. This, in turn, indicates non-invertability of the matrix, thus rendering the system insolvable, which means no solution for the Lagrangian multipliers can be found.

Now, in the modification process, if an entire row(s) or column(s) is (are) selected to be exogenously estimated, the matrix [A] will be left with all zeroes in those row(s) or column(s), which translates into insolvability of the system. In other words, under this special case the modified Lagrangian technique can not be used in its original form.

The problem, however, may be rectified via some adjustment to the system. One should start with the fact that when, for instance, an entire row is taken out, the value of the Lagrangian multiplier λ associated with that row is no longer of any value to the researcher. This is obvious since regardless of its value, λ_i will be multiplied by $\frac{x_{ij}^0}{2}$ which has zero value in the Lagrangian system (because it has been taken out and replaced by zero) and the associated estimate of the target year cell (i.e., $\tilde{x}_{ij}^t$) is also zero (again due to the fact that this value is exogenously estimated and in the Lagrangian system is treated as zero). Therefore the size of the matrix (in equation **7-49** of chapter seven) should be reduced by the number of completely known (exogenously determined) rows, and the term $\sum_j^n S_{ij}$ must be specified for $i \neq k$ (k denoting the known rows). Then the system can be solved and via the obtained values for the appropriate Lagrangian multipliers, the values of target year coefficients may be found. These estimates along with the exogenously determined values for selected rows will constitute the complete estimated matrix of the target year. In this study, the problem was relevant in the case of the MDLAGKEY variant, and the problem was removed in the manner just described.

Second, in the case of MDLAGMIP, i.e., when the most important parameters criterion suggested by Sebald (1974) and Bullard and Sebald (1975) was used, the matrix perturbation scheme employed in this study has deviated from what was suggested by the original developers of the algorithm. More specifically, in perturbing a matrix [A] by δ all elements of the matrix were multiplied by this constant. However, wherever the elements of [A] were zeros, Sebald and Bullard suggested a very small additive perturbation on these elements. This was done to assure transformation of the original zeroes to some number, hence leaving no cell in the matrix unperturbed. This procedure was not adopted in the current study. Instead, the [A] matrix was uniformly perturbed by the constant δ, which implies all original zeroes in the original matrix were preserved as zeroes after perturbation. Since modified version of RAS is also used in the current study, and since the RAS generated estimates contains zeroes in the same places as the original matrix (by virtue of the zero preservation property of the RAS procedure), for the purpose of continuity and comparison, it was deemed reasonable to drop the additive perturbation and preserve the original zeroes in the perturbed matrix when generating the estimates of 1972 table via MDLAGMIP.

Third, again in the case of MDLAGMIP, Sebald (1974) and Bullard and Sebald (1975), suggest avoiding brute force application of the Sherman-Morrison relationship. This is because the complete solution involves extremely large number of tests (n^4 to be exact, which for the (71) by (71) matrix used in this study will amount to 25,411,681 tests). To alleviate the problem they develop an algorithm that substantially reduces number of required tests by aborting the search at a given prescribed desired threshold. Considering the fact that their main objection to brute force application was due to computational limitations and the costs involved in 1974-1975, and given the fact that, thanks to powerful personal computers, these factors are no longer applicable, the most important parameters used in MDLAGMIP variant of the present study were found through performing all 25,411,681 relevant tests.

Fourth, the "adjusting coefficient" ρ of the RESMIN method (see chapter seven), chosen for the purpose of this research is (-0.10). The choice in this case was an empirical one, i.e., since the actual table was at hand, a variety of values for adjusting coefficient was examined and the value yielding the lowest mean absolute deviation (MAD) was selected. Values tested were ranged from (+1) with MAD of (0.03941) to (-1) with MAD of (0.00263) and the local minimum of MAD at (0.00203) corresponding to adjusting factor of (-0.10). Of course, this procedure is not applicable in the practical world, since the actual target year table is not available to researchers. In practice the choice of adjusting coefficient is purely subjective and based on past experiences or the best judgement of the analysts, unless a systematic algorithm for the searching purpose can be devised.

5.3) MATRIX COMPARISON AND CONCEPT OF CLOSENESS:

A crucial part of this study, as referred to in the previous section, involves the comparison of the simulated tables of the target year with the actual one in order to assess the efficiency of each projecting technique. The comparison of matrices, however, is not a straightforward task. The crux of the matter is to determine the closeness of the projected matrix to the actual survey based matrix. For instance, assume that there are two projections of a base matrix via employment of two different projecting techniques. Further assume that the first method estimated all except one of the elements of the target year matrix with hundred percent accuracy; the inaccurate element, however, is grossly estimated. The second method, on the other hand, contains no exact estimation of the elements of the target matrix, but all the estimated elements are within a given percentage of their true values. Now, the question is, which one of the two methods gives a better estimation of the target matrix?

Furthermore, one should decide that whether the closeness of estimated direct tables are more important or the accuracy of their associated Leontief inverse matrices and the respective multipliers? In other words, in matrix comparison, should researchers be more concerned with a cell-by-cell accuracy of the matrices or the goal should be a general operational accuracy of the matrices as a whole? Jensen (1980) terms these two concepts of accuracy "partitive" and "holistic" respectively, and argues that while partitive accuracy is not achievable, a holistic accuracy test is a sufficient mean of comparing matrices in an I-O analysis. The issue, however, is far from being settled. For even after deciding on the partitive or holistic accuracy, one must still choose a criterion to be used for comparative evaluation of competing techniques. A variety of accuracy measures have been proposed to address this problem. Unfortunately, no one index or criterion is viewed to be universally acceptable or superior to others in assessing the closeness in matrix comparison. This fact has been noted by many researchers such as Harrigan, et al. (1980), Kim (1984), and Israilevich (1986). Although some criterions are generally more plausible than others, the actual choice heavily depends on the objectives and needs of any given study, and in different situations one or another criterion might be more appropriate than others. The appropriateness, however, is not universal and consistent.

To remedy the situation, a "package" of complementary accuracy tests is suggested in place of a single universal criterion (e.g., Butterfield and Mules, 1980). Since each single criterion emphasizes a certain aspect of the comparison, a carefully assembled "package" of statistical tests and accuracy measures will ensure, to the extent that is possible, that the shortcomings of one criterion are compensated for through inclusion of other tests.

The present study makes use of a rather comprehensive "package" of statistical and accuracy tests. The elements of this "package" are chosen from a set of available tests, each serving a particular purpose. For reasons that will be apparent later, in this project the comparison of total outputs and multipliers are excluded and the only attempt towards a holistic comparison is application of the accuracy tests to the Leontief inverses generated through different estimation techniques. These tests are briefly explained below. The sources consulted for these explanations include Theil (1966), Pindyck and Rubinfeld (1981), Berenson and Levine (1992), Watson, et al. (1990), Lee (1993), and Griffiths, et al. (1993).

5.3.1) Wilcoxon Rank-Sum Test. The Wilcoxon rank-sum test, also referred to as the Mann-Whitney U test, is one of the accuracy tests performed in this research. This test is a non-parametric or distribution free test which utilizes the ranking of the observations in each sample, here columns of the estimated and benchmark matrices. The test can determine whether two independent samples are taken from two populations with the same relative frequency distribution, or whether two populations have the same medians. The test is a non-parametric counterpart of the t-test for equality of two means, and does not require the normality assumption.

In the context of this research, the Wilcoxon rank sum test can help to determine whether there is a consistent overestimation or underestimation of the columns of the matrices generated by various projection techniques. The null hypothesis is that, for a given level of significance, the set of estimated elements of any column in the target year matrix is not significantly above or below the corresponding set of elements of the survey based matrix. If the null hypothesis is true, then the two samples are drawn from the same population (or having the same median).

Normally the "two sample t-test" will be used to test this hypothesis. However, the t-test assumes that the two samples are randomly drawn from the populations that are normally distributed and have equal variances. The t-test also requires that the data be measured on at least an interval scale. Non-parametric tests, such as the Wilcoxon rank sum test, do not require such stringent assumptions.

The procedure first combines the two sets and ranks them. Rank one is given to the smallest value in the combined set, rank two is assigned to the second smallest value in the set, and so on until the largest number in the set assumes rank n (equal to the number of observations in the set). Designating the sum of ranks in the benchmark and projected matrices as R_1 and R_2 respectively, if the null hypothesis is true, one would expect that these sum of ranks to be approximately equal.

The Next step is to calculate the U statistics for either of the two samples and test the hypothesis. For example, the U for the first sample is obtained via:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \tag{5-1}$$

where n_1 and n_2 are number of observations in each sample, and R_1 is the total of the ranks in the first sample.

The Wilcoxon's U can be utilized to render judgment on the difference between two population distributions by measuring the difference between the ranked observations of the two samples from those populations. The U's sample distribution has the following mean and standard deviation:

$$\mu_u = \frac{(n_1)(n_2)}{2} \tag{5-2}$$

$$\sigma_u = \sqrt{\frac{(n_1)(n_2)[n_1 + n_2 + 1]}{12}} \tag{5-3}$$

For sample sizes greater than 10, the sampling distribution can be approximated by normal distribution. Thus, the standardized normal variate will be:

$$Z = \frac{U - \mu_u}{\sigma_u} \qquad (5\text{-}4)$$

The final step is to compare the calculated Z value (Z^*) with the critical Z value (Z^c) based on the desired level of significance (α). For a two tailed test,

if $Z^* < Z^c_{\alpha/2}$ or $Z^* > -Z^c_{\alpha/2}$, one will fail to reject the null

hypothesis of no significant difference between the two samples, (i.e., $H_0 : \eta_1 = \eta_2$), against the alternative hypothesis of significant difference between the two samples, (that is $H_a : \eta_1 \# \eta_2$).

Butterfield and Mules (1980) and Kim (1984) are among those who utilized the Wilcoxon rank sum test. In this test, however, positive and negative errors can offset each other. Therefore, even though large deviations might exist, the result may appear reasonable. Hence the Wilcoxon test alone will not suffice and should be complemented by other tests.

5.3.2) Regression Analysis. Another test performed in the current study is regression analysis. Regression analysis indicates the general tendencies between the two tables instead of emphasizing cell by cell comparison and can handle zero coefficients.

The general format is of the form:

$$\tilde{a}_{ij}^{t} = \alpha + \beta a_{ij}^{t} + \epsilon_{ij} \qquad (5\text{-}5)$$

where $\tilde{a}_{ij}^{t}$ and a_{ij}^{t} are projected elements of the target year and their corresponding elements in the benchmark matrix respectively. The squared correlation coefficient (R^2) and the standard F-test for a given level of significance can be used to evaluate the overall significance of the regression. The conventional t-test also may be used to test the values of the intercept (α) and the slope of the line (β). A good estimate will have a (R^2) close to one, an intercept not significantly different from zero, and a slope coefficient close to unity. The null hypothesis, then, is that for a given level of significance α and β are not significantly different from zero and unity respectively. For reasons that will be explained below, these tests are not entirely satisfactory. Therefore, in this project, in addition to the standard (F) and (t) tests, an F-statistic constructed from comparison of restricted and unrestricted least squares results have been used and the hypothesis of zero intercept and unity slope have been tested simultaneously. The logic of this test is explained below.

Results of any regression analysis are obtained via data contained in the sample. However there may be some non-sample information that can help the analysis. Then it is desirable to test the validity of these information. To achieve this, the non-sample information are imposed on the data as restrictions, and restricted estimators are obtained accordingly. If the imposed constraints are compatible with the data, their imposition should not cause too much of a reduction in the explained variation of dependent variable. If, on the other hand, the variation is significant, the imposition of restriction can be judged as unjustified. The F-statistic for testing this significance is constructed from the comparison of restricted and unrestricted estimators as:

$$F^* = \frac{\dfrac{RSS_r - RSS_{ur}}{J}}{\dfrac{RSS_{ur}}{N - k}} \qquad (5\text{-}6)$$

where RSS represents residual sum of squares or unexplained portion of the variation in the dependent variable, J is equal to the number of linear restrictions imposed based on non-sample information, N is the number of observations, k is number of parameters in the unrestricted case, and the subscripts (r) and (ur) denote restricted and unrestricted cases respectively.

The null hypothesis, then, is simultaneous validity of all (J) restrictions. In the case pertinent to the current study, the non-sample information is that the intercept of the regression line should be zero and the slope of the line should be unity. Incorporation of these information into the model will introduce two restrictions and the null hypothesis will be:

$$(5\text{-}7) \qquad H_0 : \begin{cases} \alpha = 0 \\ \beta = 1 \end{cases}$$

against the alternative hypothesis: H_1 at least one of these restrictions does not hold.

The calculated F-statistic should be compared to a critical value of F for a given level of significance and [J, (N-k)] degrees of freedom. The null hypothesis is rejected (i.e., non-sample information is not compatible with the data) if $F^* > F^c$.

Performance of this F-test will remove the difficulties associated with standard F-test of overall significance. This is because in a simple regression model, standard F-test will be identical to the t-test and the calculated F value will be equal to square of the t-value of the slope coefficient. Thus, testing the intercept and slope coefficients individually can lead to some ambiguity and a researcher, as noted among others by Round (1983), Kim (1984), and Israilevich (1986), can encounter several difficulties in ranking and comparison of various non-survey techniques via regression analysis.

For instance, if the results lead to $\alpha < 0$, $\beta > 1$ or $\alpha > 0$ and $\beta < 1$, the ranking of the projecting techniques may be difficult. The existence of some consistent pattern in the error terms, i.e., dependence of error terms, can result in a high R^2, which could indicate an accuracy in the projections that really does not exist. Moreover, the deterministic nature of I-O tables implies specification of covariance terms, hence making application of stochastic techniques difficult. Further, presence of fairly scattered observations above and below the regression line and the resultant offsetting effects of the errors can lead to the values of the intercept and the slope to be zero and one respectively, hence giving the erroneous impression of a close estimate of the target year matrix.

To somewhat alleviate the problem, a joint F-test have been proposed by Harrigan et al. (1980), and calculation of a special F-statistic was suggested. However, as noted by Israilevich (1986), if this statistic itself is not significant, then ranking of the estimation techniques from the value of this statistic will not be possible. Also, it is possible that a technique be assessed by the F-test to be superior to other methods, while other tests of closeness do not rank the same technique as high. This is due to the fact that the F-test here is more concerned with the overall similarity of two matrices than the cell by cell comparison, hence the contradictory results. Furthermore, the existence of relatively large number of zeros or very small values in an I-O table could introduce a bias in the relationship between the actual and predicted matrices.

For these reasons, then, the regression analysis alone may not yield reliable results with regards to absolute and relative accuracy of various estimation techniques, and it should be used in conjunction with other complementary tests. It should be noted, however, that much useful information can be obtained through usage of regression analysis, and the results readily lend themselves to tests of significance.

It is the recognition of these merits that led to the utilization of regression analysis by many researchers such as Schaffer and Chu (1969), Morrison and Smith (1974), Harrigan et al. (1980), Butterfield and Mules (1980), Kim (1984), and Israilevich (1986).

5.3.3) __Chi-Square Test__. The third statistical test used for evaluation of the estimated and actual matrices is the chi-square goodness of fit test. Generally, this test compares the actual values in each category to the values that theoretically would be expected to occur if the data followed the selected probability distribution, and it is a measure of absence of compatibility between the sample data and the hypothesis. The statistic gathers together the discrepancies between the observed and expected values and the null hypothesis tests whether the calculated chi-square is reasonable for a given degree of significance or not. Specifically, in matrix comparison test, the chi-square goodness of fit test compares the estimated and the actual

I-O coefficients. The null hypothesis is whether the distributions of $\tilde{a}_{ij}^{t}$ and a_{ij}^{t} are the same, against the alternative hypothesis that they are not.

The chi-square statistics for the two set of data (estimated and benchmark matrices in this case) is found through:

$$\chi^2 = \sum \frac{(\tilde{a}_{ij}^{t} - a_{ij}^{t})^2}{a_{ij}^{t}} \qquad (5\text{-}8)$$

where $\tilde{a}_{ij}^{t}$ and a_{ij}^{t} are the estimated and actual values of the target year matrix respectively. The null hypothesis can be rejected if the calculated χ^{*2} is greater than a critical value for a specified desired level of significance. (i.e., $\chi^{*2} > \chi_c^2$). Here, a good estimate will result in a small value of the chi-square. For if the size distribution of the estimated and actual matrices is similar, the errors will be small, hence a small value for the chi-square statistic.

As noted by Harrigan et al. (1980), Morison and Smith (1974), and Kim (1984), there could be some problems with the chi-square test. Most notably, when $\tilde{a}_{ij}^{t}$ is not zero but a_{ij}^{t} has a zero value. The aforementioned researchers suggest either omission of these zero and non-zero pairs, or assignment of extremely small value to each zero a_{ij}^{t} in order to avoid division by zeroes. But the chi-square statistic is very sensitive to a small number in the denominator and this remedy might introduce a bias into the calculation. Furthermore, as noted by Israilevich (1986) the omission of some of the elements from calculation of the χ^2 statistic may lead to misleading conclusions. In any event, as suggested by some analysts [e.g., Kim (1984)], it is best to use this statistics as a relative measure of distance and not absolute goodness of fit, and perform the comparisons over size distribution of the coefficients.

The chi-square test has been utilized by researchers such as Morrison and Smith (1974), Harrigan et al. (1980), Butterfield and Mules (1980), and Kim (1984). In the current project this test is performed over forty classes as well as the entire tables without classifications.

In passing, it should be noted that this test is greatly affected by classifications and the classes with expected zeros. This in turn, reduces the power of the chi-square test substantially. An elaboration of these points will be offered in the next chapter.

5.3.4) **Mean Absolute Deviation**. Another statistic used in the comparison of the projections in this study is the mean absolute deviation (MAD). This statistic is a summary measure and shows the distance between the actual and estimated data, i.e., the average amount of difference between an estimated coefficient and its corresponding true value in the benchmark table. For two (n) by (n) matrices, it simply averages the elements in the error matrix (E) without any regard to their signs. (E) is the matrix constructed from the difference between the estimated matrix and the benchmark matrix. Formally:

(5-9) $$E = \tilde{A}_{ij}^{t} - A_{ij}^{t}$$

and

(5-10) $$MAD = \left(\frac{1}{N^2}\right) \sum_{i=1}^{n} \sum_{j=1}^{n} |e_{ij}|$$

This statistic is affected by presence of large errors, thus it should be considered along with the mean and standard deviation of the coefficients in order to evaluate the performance of each projection technique. The value of MAD per se can not be evaluated and statistical tests of significance are not feasible. It can be used only in ranking of the estimation techniques in terms of their accuracy. The lower the value of MAD associated with an estimation method, the better is the estimate. A zero value of MAD indicates a perfect estimate.

Most researchers include this statistic in their analysis due to its ease of calculation and interpretation. Among these, Morrison and Smith (1974), Hinojosa (1978), Harrigan et al. (1980), Butterfield and Mules (1980), Sawyer and Miller (1983), Kim (1984), and Israilevich (1986) Afrasiabi and Casler (1991) might be mentioned.

5.3.5) Coefficient Of Equality. A statistic designed for measurement of closeness of two matrices and defined as the ratio of the projected value of each element of the updated matrix to its corresponding value in the actual survey based matrix, i.e.,:

$$\theta = \frac{\tilde{a}_{ij}^{t}}{a_{ij}^{t}} \tag{5-11}$$

It is evident that the closer θ to unity the better is the estimate. To carry out the test in a reasonable fashion, the data should be divided into several classes and degree of approximation (1- θ) be calculated. Then by analyzing the values for degree of approximation, it is possible to determine the number of coefficients that are estimated within a predetermined percentage of their true values. This, along with the mean, standard deviation, maximum value, and the distribution of θ will provide some insight into the performances of the alternative estimation techniques and aid the researcher in evaluation and ranking of these techniques.

The coefficient of equality also has some deficiencies as it will not be of much use in cases when both $\tilde{a}_{ij}^{t}$ and a_{ij}^{t} are zeros or when one is zero and the other has very small value. This statistic is used by Kaneko (1983), and has been utilized here.

5.3.6) <u>Standardized Total Percentage Error</u>. The standardized total percentage error (STPE) is another statistic used as a measure of distance between two sets of data (two matrices in this case). It expresses the average of absolute discrepancies among estimated and actual data sets as percentage of the mean of actual data. Hence, its value may be used as an average measure of distance between two sets of data expressed as percentage deviation from the mean of the base data. It can be used in ranking of the estimation methods, but its value in itself can not be evaluated and tests of significance are not available. The smaller the value of STPE, the better is the estimate, and a STPE equal to zero indicates a perfect estimate.

This statistic is calculated through dividing the mean absolute deviation of elements of the estimated and benchmark matrices by the mean of elements of the benchmark matrix. Formally this may be expressed as:

$$(5\text{-}12) \qquad STPE = \left[\frac{\left(\frac{1}{n^2}\right)\sum_{i=1}^{n}\sum_{j=1}^{n} |\tilde{a}_{ij}^{t} - a_{ij}^{t}|}{\left(\frac{1}{n^2}\right)\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^{t}} \right] \times (100)$$

Or more compactly:

$$(5\text{-}13) \qquad STPE = \left[\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} |\tilde{a}_{ij}^{t} - a_{ij}^{t}|}{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^{t}} \right] \times (100)$$

This measure is also very sensitive to zero or small values of a_{ij}^{t} . Miller and Blair, as reported by Israilevich (1986), strongly advocate the usage of this measurement, and researchers such as Afrasiabi and Casler (1991), utilize it.

5.3.7) **<u>Root Mean Square</u>.** In comparison of an actual and simulated sets of data, a desirable measure to a researcher is a measure enabling him/her to examine how closely the individual estimated variables resemble their actual values in the original population. Root mean square of error terms or RMS, that is root mean square of elements of the error matrix E (a matrix constructed from the difference between the actual and estimated table), is one such measure.

The comparison of estimated and actual matrices fits well within the applications of RMS, making it a popular measure in matrix comparison. Thusly, it has been utilized here. It has been also used by analysts such as Casler and Afrasiabi (1991).

Formally, RMS for comparison of two (n) by (n) matrices can be defined as:

(5-14)
$$RMS = \sqrt{\frac{\sum_{i}^{n}\sum_{j}^{n} e_{ij}^{2}}{N^{2}}}$$

where $e_{ij} = \tilde{a}_{ij}^{t} - a_{ij}^{t}$, i.e., the difference between the individual estimated and actual values, or elements of the error matrix E.

To render a judgment about the estimated data, the value of RMS should be compared with the average of the variable under study. In ranking of the updating techniques, however, the lower the value of RMS, the better is the method, since the point of reference for all of them is the same. A zero value for RMS indicates a perfect fit. It should be noted that RMS penalizes more heavily those estimation techniques that generate large individual errors.

5.3.8) **Theil's U**. Theil's inequality coefficient or Theil's U is another statistic used in comparison of a simulated and an actual set of data. It is well suited for comparison of estimated and actual I-O matrices and is widely used (e.g., Afrasiabi and Casler, 1991), hence included in the criterions employed in this endeavor. The U statistic utilizes the RMS, and it is constructed in such a way that its value always falls between zero and one. The closer its value to zero, the better is the estimate. U equal to zero indicates a perfect fit and U equal to one signifies the worst possible case.

Formally Theil's U for comparison of two (n) by (n) matrices can be stated as:

(5-15)
$$U = \frac{\sqrt{\dfrac{\sum_i \sum_j (\tilde{a}_{ij}^t - a_{ij}^t)^2}{N^2}}}{\sqrt{\dfrac{\sum_i \sum_j (\tilde{a}_{ij}^t)^2}{N^2}} + \sqrt{\dfrac{\sum_i \sum_j (a_{ij}^t)^2}{N^2}}}$$

It should be noted that the numerator is nothing but the RMS error between the estimated and actual matrices.

The Theil's U can be decomposed into three different components, namely UM, US, and UC. These component parts, in turn, allude to the sources of the estimation error and will furnish the researchers with some additional insights in error analysis and data comparisons. This decomposition has been carried out in the current study and the results, along with the results of other criterions are reported in the following chapters. A brief description of the components of the Theil's U should suffice here.

5.3.9) UM. This statistic, representing the bias or systematic error, measures the extent of deviation of the average values of estimated and actual data sets. Its magnitude can vary between zero and one, and is directly related to existence of systematic error in the estimation. The lower the U statistic, the lower is the systematic bias in the estimated set. Value of zero is an indication of an estimate, free of systematic error.

The UM statistic can be calculated by dividing the square of the difference between the means of estimated and actual variables by the square of RMS.

For comparison of two (n) by (n) matrices it may be stated as:

$$UM = \frac{(\tilde{\mu} - \mu)^2}{\frac{\sum_i \sum_j (\tilde{a}_{ij}^t - a_{ij}^t)^2}{N^2}} \qquad (5\text{-}16)$$

where $\tilde{a}_{ij}^t$ and a_{ij}^t represent the estimated and actual elements of the matrices in question, with $\tilde{\mu}$ and μ denoting the means of the two series respectively.

5.3.10) US. The variance part of the Theil's U and is denoted as US. Its magnitude can fluctuate between zero and one, and is an indication of existing inefficiency of measurement of the degree of association in variability of the actual and estimated data. The larger the value of US, the higher the inefficiency of estimates (or the lower the degree of association between the two data sets). US equal to one indicates wide fluctuation in one set with no change in the other. For two (n) by (n) matrices, this statistic may be calculated via:

$$US = \frac{(\tilde{\sigma} - \sigma)^2}{\frac{\sum_i \sum_j (\tilde{a}_{ij}^t - a_{ij}^t)^2}{N^2}} \tag{5-17}$$

where $\tilde{\sigma}$ and σ represent the standard deviations of the $\tilde{a}_{ij}^t$ and a_{ij}^t series respectively, and the denominator is square of the RMS.

5.3.11) UC. The last component of the Theil's U is the covariance part. It accounts for the balance of errors in the estimate, after compensation for bias (systematic error or deviations from the average values) and inefficiency.

The value of this measure can fluctuate from zero to one, and may be found via:

(5-18)

$$UC = \frac{2(1-\rho)(\tilde{\sigma})(\sigma)}{\frac{\sum_i \sum_j (\tilde{a}_{ij}^t - a_{ij}^t)^2}{N^2}}$$

where ρ represents the correlation coefficient of the two data series and other terms are as before. Knowing that the total "inequality coefficient" or Theil's U is equal to one, the UC also can be simply obtained through subtraction of bias and inefficiency components from one, i.e.,:

(5-19) $$UC = 1-(UM + US)$$

For any value of $U > 0$, the higher the UC the better the estimate, with UC equal to one representing the best case, since it makes the errors due to bias and inefficiency equal to zero.

5.3.12) **Mean, Standard Deviation, and Maximum Value**. In addition to the previously mentioned tests, for each estimated matrix, the mean and standard deviation along with maximum values of the estimated coefficients are also calculated and reported. These statistics should provide some basis for comparison of actual data and various estimates, by indicating the measures of central tendency and dispersion for the actual data and estimated tables. The standard formulas for mean and standard deviations for case of (n) by (n) matrix are utilized here, i.e.:

(5-20) $$\tilde{\mu} = \frac{\sum_i \sum_j \tilde{a}_{ij}^t}{N^2}$$

and

(5-21) $$\tilde{\sigma} = \sqrt{\frac{\sum_i \sum_j (\tilde{a}_{ij}^t - \tilde{\mu})^2}{N^2}}$$

where all terms are as previously specified.

As noted before, there are many other methods suggested for measuring the closeness of two matrices. A selected set of these methods is chosen to be utilized in this study. Many other such tests of matrix comparisons are available but excluded in the current project. A partial list of these tests includes mean absolute percentage error or MAPE, used by Czamanski and Malizia (1969), Miller and Blair, (1985); standardized mean absolute deviation or SMAD, used by Butterfield and Mules (1980) and Kim (1984); mean similarity index or MSI, used by Morison and Smith (1974), Harrigan et al. (1980), and Kim (1984); contingency analysis used by Israilevich (1986); information index used by Czamanski and Malizia (1969), Morrison and Smith (1974), and Harrigan et al. (1980); and Euclidian difference used by Harrigan et al. (1980).

All suggested methods suffer from some deficiency and none of them can be chosen singularly to conduct a satisfactory test of closeness for matrices. The exclusion of some of these tests in the current study is due to the fact that a rather comprehensive, complementary, and all encompassing package of tests is utilized and inclusion of additional tests neither yields any new information about the matrices nor remedies any of deficiencies present in the current package. Furthermore, for zero coefficients, some of the excluded tests will involve division by zero, which makes those particular statistics inapplicable for dissaggregated tables, hence their exclusion in the current undertaking. Lastly, as noted by Round (1983), comparison of various updated tables with the actual table generally do not present much information about the absolute efficiency of the respective techniques. Therefore, these tests should really be used as a relative measure of efficiency of various updating techniques.

The estimated matrices obtained via application of the selected updating techniques were subjected to the above mentioned package of closeness tests and the results of these comparisons are reported in the next chapter.

CHAPTER SIX

RESULTS

"Results! Why, man, I have gotten a lot of results. I know several thousand things that won't work."

Thomas Edison

Utilizing the selected updating methods, nine estimated matrices for both direct and inverse coefficients were obtained. The computational demand of none of the techniques was heavy and, using an IBM PC 486 DX2, in no case required more than ten minutes of computer time. In the cases involving iterative solutions, i.e., RAS and all its derivatives, convergence always were arrived with less than 90 iterations.

6.1) INTRODUCTION:

In order to evaluate the performance of each projecting technique as well as determining their rankings, the estimated tables were compared to the benchmark table. The comparison involved utilization of a "package" of statistical tests. Details of the tests used in the current project were given in chapter five and the "package" itself included Wilcoxon rank-sum test, regression analysis, chi-square statistics, mean absolute deviation, coefficient of equality, standard total percentage error, root mean square, and Theil's U along with its components UM, US, and UC.

As mentioned before, the comparison could be for either cell-by cell (partitive) or operational (holistic) accuracy. Generally speaking, the comparison can be conducted at four different levels. Moving from partitive to holistic comparisons these levels are:

a) cell-by cell comparison of the technology matrices, which is the most partitive one. In this comparison the individual estimated coefficients are compared to their counterparts in the benchmark table, and the estimation techniques are evaluated accordingly. The emphasis here is on individual cell's accuracy and direct effect as oppose to the total aggregate accuracy.

b) cell-by-cell comparison of the Leontief inverse matrices which is a step closer to holistic comparison. For each element of this matrix represents both direct and indirect effects of one unit change in final demand of any given industry on total output of that industry. In this case accuracy of each individual cell is concerned, but this concern is not limited to the direct effect alone and it encompasses indirect effects as well. In other words the estimated tables (hence the estimation techniques) are evaluated on the basis of their ability to produce accurate means of measuring the direct and indirect effects of changes in the final demand.

c) comparison of output multipliers, which is the next step towards holistic comparison. This is due to the fact that output multipliers represent the total direct and indirect effects of one unit change in the final demand of a given industry's output throughout the whole economy. In this case, accuracy of each individual cell is not of main concern. Instead, the accuracy of the overall aggregate effect of a change in final demand of a given industry throughout the economy is emphasized. The techniques, then, are evaluated based on their performances in generating accurate multipliers as opposed to accurate individual cells.

d) Comparison of total outputs is the last stage and most holistic one. In this case the accuracy of total output is the main issue. That is to say a technique that can estimate the total output closer to the real value is considered to be the superior technique without regards to the accuracy of individual cells generated by that technique.

In the current study, the first two comparisons are performed, but the third and forth one are left out. The basic logic behind this choice is derived from the nature of the chosen estimation techniques and the objectives of this study. The main purpose of this research is evaluation and comparison of non-survey techniques. Comparison of the technology and the Leontief inverse matrices generated by the chosen techniques can lead to evaluation and ranking of these methods. Specifically speaking, via comparison of the estimated technology and the Leontief inverse matrices with their benchmark counterparts, it may be possible to evaluate the relative performance of the estimation techniques and determine the goodness of these estimates. Comparison of the multipliers and total outputs, on the other hand, in the case of estimation techniques utilized in the present project, will not yield any substantial information about these methods. This is due to the nature of these techniques and the way their objective functions are set up. For all these techniques will produce a total output that is hundred percent accurate and multipliers that are extremely close to the benchmark values, hence implying an accurate estimation even though there could be drastic variations at the cell-by-cell or industry-by-industry levels. In other words the multiplier and total output comparisons inevitably will yield satisfactory results due to the way the techniques are constructed, hence rendering the comparisons unnecessary. Thus, the comparisons in this project are limited to direct and inverse coefficients only.

6.2) EVALUATION OF THE RESULTS:

At this juncture, attention can be directed towards the two sets of comparisons for the estimated direct and inverse matrices. Each criterion is applied to direct and inverse estimated matrices and the results are reported and compared in the reminder of this chapter.

6.2.1) Number of Negative Coefficients. As was mentioned earlier, not all of the estimation methods guarantee generation of non-negative coefficients. Hence, as expected, there are some negative coefficients in the estimated matrices. Generally, if the number of negative coefficients generated via an estimation technique are small relative to the size of total estimated coefficients, the problem can be circumvented by either disregarding them or by substituting zeros in their places. The latter approach is particularly appropriate if the negative coefficients in the estimated matrix correspond to zero entries in the base matrix. Table 6-1 below presents a summary of negative

TABLE 6-1

NUMBER OF NEGATIVE COEFFICIENTS GENERATED BY EACH TECHNIQUE

TECHNIQUE	DIRECT	INVERSE
NAIVE	0	0
RAS	0	0
RECRAS	2	0
PROPVA	0	0
FRIED	2	0
RECLAG	2	0
RASLAG	0	0
RERALA	2	0
ALMON	2103	1708

coefficients, both for direct and inverse matrices, generated by various estimation techniques in this research. Four of these techniques, i.e., NAIVE, RAS, PROPVA, and RASLAG, by virtue of their properties can not generate negative coefficients. The results, as shown in table 6-1, are in accordance with *a priori* expectation for these techniques. Four out of the remaining five methods, namely RECRAS, FRIED, RECLAG, and RERALA, did not yield any negative coefficients in the inverse matrix cases, while the estimated matrices generated by them for direct coefficients contained two negative values. The number of negative coefficients (two), when compared to the total number of coefficients generated by each method (5041), is negligible and do not pose a major problem in evaluation of the techniques, hence they can safely be discarded or be replaced by zeros. The last method, the Almon approach, however, is quite different. For it generated 2103 negative coefficients (41.7% of all estimated coefficients) in the direct matrix and 1708 negative coefficients (33.9% of all estimated coefficients) in the inverse matrix. These negative coefficients can neither be disregarded nor one can utilize zero substitution remedy to deal with them. The experiment here clearly demonstrates that the Almon's minimand is far less robust than other techniques in assuring generation of non- negative coefficients.

6.2.2) Non-Parametric Test. The Wilcoxon rank-sum test was conducted on column by column basis of the estimated direct and inverse matrices with those of the 1972 benchmark table to investigate the existence of any consistent over or under estimation of the columns. The null hypothesis of equality of the medians of each estimated column and its counterpart in the benchmark table- which suggests no significant over or under estimation of columns- was tested at 5% level of significance against the alternative hypothesis of significant difference between the two sets.

Formally:

$$H_0 : \eta_1 = \eta_2 \qquad (6\text{-}1)$$
$$H_a : \eta_1 \neq \eta_2$$

for 5% level of significance, which for a two tailed test implies $Z^c = \pm 1.96$.

For direct matrices, in the cases of NAIVE, RECRAS, PROPVA, and RERALA methods, one will fail to reject the null hypothesis in all but one column. Hence it can be inferred that no column was overestimated by these methods and only one column was underestimated. For RAS, FRIED, and RASLAG methods the null hypothesis can be rejected for two columns, suggesting two columns are overestimated or underestimated by these methods. It is interesting to note that overestimated and underestimated columns under all methods were the same columns, which could be due to some structural change in those particular industries during 1966-1971 period in the Soviet Union. RECLAG and Almon approach overestimated zero and sixteen columns and underestimated ten and forty two columns respectively.

In the case of inverse matrices, the null hypothesis can not be rejected for the results generated by five techniques i.e., RAS, RECRAS, FRIED, RASLAG, and RERALA. Hence it can be concluded that these techniques consistently neither overestimated nor underestimated the columns of inverse matrices. NAIVE, PROPVA, and RECLAG underestimated the columns in one, seven, and fifteen instances respectively, with no case of overestimation. The ALMON method underestimated 44 and overestimated two columns.

The results can be visually verified via examination of figures A-1 through A-18 of the Appendix (A). Based on these results it may be stated that all techniques except RECLAG and ALMON perform satisfactorily in the case of direct coefficient matrices. For inverse matrices, the results are acceptable for all methods except RECLAG, PROPVA, and ALMON.

The ranking of the models based on number of over or under estimation of columns, as the acceptance interval narrows, may be listed as follow. For the case of direct coefficients, RAS & RASLAG, FRIED & NAIVE, RECRAS & PROPVA, RERALA, RECLAG, and ALMON. For the inverse coefficients the ranks are RAS & FRIED & RASLAG, RERALA, RECRAS, NAIVE, PROPVA, RECLAG, and ALMON.

6.2.3) Coefficient of equality. This statistic measures the degree of closeness of the estimated coefficients to their actual values. The closer the value of the coefficient of equality to unity, the better the estimate.

Table 6-2 summarizes the results of this test for various techniques. The statistics for actual table is also included here to facilitate the comparison. The "degree of approximation" (1 - coefficient of equality) is used to

TABLE 6-2

SUMMARY OF ESTIMATING PERFORMANCES, DIRECT COEFFICIENTS

ESTIMATION METHOD	DEGREE OF APPROXIMATION (1- COEFF. OF EQUALITY)			COEFFICIENT OF EQUALITY		
	5%	10%	20%	MEAN	SD	MAX
ACTUAL	5041	5041	5041	0.8972	0.3037	1
NAIVE	783	1044	1579	2.3107	23.032	1073.40
RAS	820	1083	1643	2.0403	16.260	661.82
RECRAS	805	1030	1575	2.0147	16.011	611.97
PROPVA	788	1069	1567	2.0553	19.896	866.77
FRIED	792	1055	1599	2.0183	16.066	683.47
RECLAG	677	828	1147	1.3717	11.595	452.86
RASLAG	820	1083	1643	2.0403	16.260	661.82
RERALA	801	1038	1571	2.0143	15.990	611.96
ALMON	593	677	836	112.78	3407.2	1093700

determine the total number of estimated coefficients generated by each method which lie within 5%, 10%, and 20% error of their true values.

Accordingly, it can be seen that at 5% error, the NAIVE method yields results which are superior to those of RECLAG and ALMON. Out of those methods that outperform the NAIVE method, the ranking based on prediction of largest number of coefficients within 5% of true value is RAS, RASLAG, RECRAS, RERALA, FRIED, PROPVA, and RECLAG. At 10% error, NAIVE method surpasses RECRAS, RECLAG, RERALA, and ALMON. The ranking among the remaining methods is RAS, RASLAG, PROPVA, and FRIED. At 20% error level, the NAIVE approach outperforms RECRAS, PROPVA, RECLAG, RERALA, and ALMON. The ranking among the other methods is RAS, RASLAG, and FRIED.

It is evident from the table 6-2 that RAS and RASLAG methods yield identical results that is superior to all other techniques. However, RAS method attains these results faster and with lesser amount of resources. Thus, it is the more efficient of the two. The place of other techniques in the ranking is not stable and varies as the degree of accuracy requirements changes.

It is worthy to note that as the tolerance for error is increased the NAIVE method present itself as a more viable approach. This in turn could be an indication of relative stability of large block of coefficients in an I-O table.

Table 6-2 also contains the mean and standard deviation, as well as the maximum value of the coefficients of equality associated with each technique. These statistics can provide some insight into the distribution of the coefficients estimated by various techniques. The ideal situation, i.e., hundred percent accurate estimates, the mean value will be equal to one, with standard deviation and maximum value being zero and one respectively. It is clear from table 6-2 that none of the techniques used here, satisfy these conditions. Means of the coefficients of equality range from 1.3717 to 112.78, and the standard deviations vary from 11.595 to 3407.2. Maximum values of the coefficients of equality range from 452.86 to 1093700. The ranking based on the means of coefficients of equality is RECLAG, RERALA, RECRAS, FRIED, RAS & RASLAG, PROPVA, NAIVE, and ALMON. The ranking based on degree of overestimation (the maximum values of the coefficients of equality) is RECLAG, RERALA, RECRAS, RAS & RASLAG, FRIED, PROPVA, NAIVE, and ALMON.

Therefore, based on this test, it seems that even though some of the techniques can estimate a fair number of individual coefficients within an acceptable interval, none is capable of limiting the variation in the individual estimated coefficients. The apparent poor performances, more than being an evidence of the estimating power of the techniques, may be due to existence of some very small coefficients in the actual data which in turn can create very large values for the standard deviation and maximum values of the coefficients of equality.

In the case of inverse coefficients, as it can be seen from table 6-3, the NAIVE model estimates 10.7% of the coefficients within 5% of their true values. If wider intervals of 10% and 20% are to be used, the estimated

TABLE 6-3

SUMMARY OF ESTIMATION PERFORMANCES, INVERSE COEFFICIENTS

ESTIMATION METHOD	DEGREE OF APPROXIMATION (1- COEFF. OF EQUALITY)			COEFFICIENT OF EQUALITY		
	5%	10%	20%	MEAN	SD	MAX
ACTUAL	5041	5041	5041	1.0000	0.0000	1
NAIVE	544	990	1873	1.0205	1.0325	48.276
RAS	740	1389	2512	1.0713	.79830	32.374
RECRAS	590	1107	2006	1.0473	.81329	28.379
PROPVA	399	759	1443	.86392	.91304	44.835
FRIED	730	1367	2463	1.0772	.82197	35.253
RECLAG	309	581	1212	.76614	.69128	30.067
RASLAG	740	1389	2512	1.0713	.79830	32.374
RERALA	674	1203	2200	1.0702	.80280	27.621
ALMON	161	311	586	-53.968	913.13	652.96

coefficients within those intervals by the NAIVE model increase to 19.6% and 37% respectively. These results are superior to those obtained by PROPVA, RECLAG, and ALMON models, suggesting that when inverse matrices are concerned, using the coefficients of the most recent actual table is better than utilization of any of these estimation techniques. Among the remaining models, RAS and RASLAG are tied for the first place by estimating 14.7%, 27.6%, and 49.8% of the coefficients under 5%, 10%, and 20% error intervals. RAS approach, however, is the more efficient one due to its lesser requirement. The ranking of the remaining techniques are FRIED, RERALA, and RECRAS. It is noteworthy that the ranking of all models in accurately estimating of the coefficients of the inverse matrix does not change, as the accuracy interval widens.

Tables B-1 through B-18 of the Appendix B show the frequency distribution of the coefficients of equality for both direct and inverse coefficients, and figures E-1 through E-18 of the Appendix E graphically depict these distributions.

Table 6-3 also reveals that, except in the case of ALMON model, the means of the coefficients of equality for the inverse matrices are sufficiently close to their ideal value of one. The relevant standard deviations, however, are high and suggest a high variation in the estimates. The maximum values of the coefficients of equality range from low of 27.621 to high of 652.96. These values, although less than their counter parts in the direct tables, further reaffirm the existence of wide range of estimates.

The ranking of the estimations based on the means of the equation of equality is NAIVE, RECRAS, RERALA, RAS & RASLAG, FRIED, PROPVA, RECLAG, and ALMON. Based on the standard deviations, the ranking will be RECLAG, RAS & RASLAG, RERALA, RECRAS, FRIED, PROPVA, NAIVE, and ALMON. The maximum values criterion ranks the models as RERALA, RECRAS, RECLAG, RAS & RASLAG, FRIED, PROPVA, NAIVE, and ALMON.

6.2.4) __Regression analysis__. Regressions of the direct and inverse coefficients estimated via various techniques on the benchmark coefficients were conducted as another step of accuracy tests. The detail results are presented in tables D-1 through D-18 of Appendix D and summary results associated with these tables are provided in tables 6-4 and 6-5 below. As noted before, when the constant (α) is zero and the slope (β)is equal to one, a perfect fit exists.

In the case of direct coefficients, all methods except Almon exhibit a high R^2 , indicating the existence of significant explanatory power of the regression line. The low value of R^2 in Almon model, given the fact that its corresponding regression line is statistically significant, could be due to the large variation in the estimated values of coefficients.

The Durbin-Watson test for existence of autocorrelation at 1% and 5% level of significance is performed and, except in the case of Almon, one will fail to reject the null hypothesis that error terms are not correlated. Thus, absence of serial correlation among the error terms and

TABLE 6-4

SUMMARY RESULTS OF REGRESSION ANALYSIS

Direct Coefficients

Method	Intercept	t-value	Slope	t-value	R^2
NAIVE	-0.000172	-1.349	0.982869	-4.78	0.937
	(.000128)		(.003605)		
RAS	0.000039	0.386	0.994837	-1.82	0.960
	(.000101)		(.002845)		
RECRAS	-0.000159	-1.273	1.000522	0.15	0.941
	(.000125)		(.003523)		
PROPVA	0.000167	1.270	0.868193	-35.55	0.916
	(.000131)		(.003708)		
FRIED	0.000027	0.259	0.996395	-1.22	0.957
	(.000105)		(.002967)		
RECLAG	-0.001343	-8.882	1.058289	13.67	0.924
	(.000151)		(.004263)		
RASLAG	0.000039	0.386	0.994837	-1.82	0.960
	(.000101)		(.002845)		
RERALA	-0.000177	-1.456	1.010109	2.94	0.945
	(.000122)		(.003436)		
Almon	0.000430	0.403	0.730842	-8.93	0.104
	(.001069)		(.030157)		

* Figures in parentheses are standard errors.

TABLE 6-5

SUMMARY RESULTS OF REGRESSION ANALYSIS

Inverse Coefficients

Method	Intercept	t-value	Slope	t-value	R^2
NAIVE	-0.000916	-4.492	0.981576	-13.24	0.990
	(.000204)		(.001391)		
RAS	0.000047	0.316	0.998317	-1.65	0.995
	(.000150)		(.001021)		
RECRAS	-0.000298	-1.342	0.992403	-5.03	0.988
	(.000222)		(.001512)		
PROPVA	-0.002012	-7.540	0.960607	-21.66	0.982
	(.000267)		(.001819)		
FRIED	0.000096	0.615	0.996596	-3.21	0.994
	(.000156)		(.001061)		
RECLAG	-0.002575	-9.048	1.035348	18.22	0.983
	(.000285)		(.001940)		
RASLAG	0.000047	0.316	0.998317	-1.65	0.995
	(.000150)		(.001021)		
RERALA	-0.000211	-1.126	1.002572	2.02	0.992
	(.000187)		(.001274)		
Almon	-0.002328	-3.584	0.879497	-27.21	0.887
	(.000650)		(.004428)		

* Figures in parentheses are standard errors.

strong relationship among the estimated and actual coefficients may be inferred.

Regression equations for the inverse coefficients also show high R^2, which indicates high degree of explanatory power between actual and estimated inverse matrices.

The Durbin-Watson test at 1% and 5% level of significance for existence of serial correlation is conducted. The results reveal that at 1% level the null hypothesis can not be rejected in all cases except ALMON. That is, with the exception of ALMON method, none of the methods exhibit correlations among the error terms. At 5% level, NAIVE, RECRAS, PROPVA, and ALMON suggest existence of autocorrelation, while other do no suffer from this problem.

The standard F-test at 1% and 0.5% levels of significance, which implies corresponding critical values of F to be F_c=6.63 and 7.88 respectively, is also conducted for overall significance of the regression equations. Based on these results, which can be verified from tables F-1 through F-18 of appendix F, all regression equations for both direct and inverse coefficients are statistically significant, i.e., their calculated F-values far exceed the critical values of 6.63 and 7.88 and in all cases one will reject the null hypothesis of no linear relationship among estimated values of coefficients and their actual counterparts in favor of the alternative hypothesis of existence of linear relationships.

The rankings of estimation techniques based on R^2 and overall F-test at 1% and 0.5% are identical as follows, RAS & RASLAG, FRIED, RERALA, RECRAS, NAIVE, RECLAG, PROPVA, and ALMON for direct and RAS & RASLAG, FRIED, RERALA, NAIVE, RECRAS, RECLAG, PROPVA, and ALMON for inverse coefficients.

The null hypothesis of $\alpha = 0$ was tested through usage of t-test, using 5% and 1% level of significance (critical values of 1.96 and 2.58 respectively). For the direct coefficients, as it is clear from the table 6-4, in all cases except RECLAG, one will fail to reject the null hypothesis. This result indicates that all techniques, except RECLAG, produced value of intercept that is, in accordance with a *priori* expectation of statistically not being different from zero. For the inverse coefficients, as can be seen from table 6-5, one will fail to reject the null hypothesis of zero intercept in all cases but NAIVE, RECLAG, PROPVA, and ALMON. That is to say, estimated inverse coefficients via these techniques may not good estimates. The rankings based on the intercept's t-values are FRIED, RAS & RASLAG, ALMON, PROPVA, RECRAS, NAIVE, RERALA, and RECLAG for direct and RAS & RASLAG, FRIED, RERALA, RECRAS, ALMON, NAIVE, PROPVA, and RECLAG for inverse coefficients.

In the case of slope coefficients (β), however, the calculated t-value can not directly be taken from the results of regression analysis. Instead, the appropriate t-values must be calculated from the given information. The reason for this necessity lies in the way that most statistical packages are designed. Normally, in regression analysis, one is concerned with the validity of inclusion of a given independent variable in the model. Hence, it is desired to test the significance of each variable by testing whether the estimated values are significantly different from zero or not. In other words, the null hypothesis of $\beta = 0$ is evaluated via usage of the t-test, and the calculated t- values given by statistical packages are constructed with this in mind.

In the present study, however, the hypothesized value of β is one and not zero. That is to say the null hypothesis is whether β equals to one or not. Therefore, the t-values must be calculated accordingly and then be compared with the corresponding desired critical values. The t-statistics reported in tables 6-4 and 6-5 are calculated in this manner via employment of the following relationship:

$$t^{*} = \frac{b - \beta}{\sigma_b} \tag{6-2}$$

where b and σ_b are respectively the slope coefficient and its standard deviation obtained via the regression line, and β is the hypothesized value of the slope coefficient (one in this case).

Examination of table 6-4 reveals that for direct coefficients one will reject the null hypothesis in case of

NAIVE, PROPVA, RECLAG, RERALA, and ALMON at both 5% and 1% levels (albeit barely for RERALA at 5% level). In other words, estimates obtained via these methods do not correspond to the *a priori* expectation of $\beta = 1$, hence the estimates may not be judged as good estimates. In the case of remaining techniques, since the null hypothesis can not be rejected, it may be concluded that the value of intercept coefficient is not statistically different from one, thus good estimates are obtained.

Table 6-5 can be used in the same manner for the inverse coefficients. Accordingly, it can be seen that at 1% level the null hypothesis of $\beta = 1$ is rejected for NAIVE, RECRAS, PROPVA, RECLAG, FRIED, and ALMON. For the remaining estimates the null hypothesis can not be rejected. Thus the slope is not statistically different from one, indicating a good fit. At 5% level the null hypothesis is rejected for NAIVE, RECRAS, PROPVA, FRIED, RECLAG, ALMON, and (barely) RERALA, suggesting the estimates are not good. For other techniques, one will fail to reject the null hypothesis, signifying a good fit for the regression equations.

The rankings based on slope's t-values are RECRAS, FRIED, RAS & RASLAG, RERALA, NAIVE, ALMON, RECLAG, and PROPVA for direct and RAS & RASLAG, RERALA, FRIED, RECRAS, NAIVE, RECLAG, PROPVA, and ALMON for inverse coefficients.

Two clarifying notes, however, are in order. First, in the case of overall significance of simple regression, the standard F-test is not really a meaningful test. For the standard F-test is test of joint slopes, and when $\alpha = 0$ it will loose its meaning and will reduce to the t-test for the slope coefficient (the F-values are squares of the slope coefficients' t-values). Hence a more suitable F-test is desired. Second, individual t-tests for slope and intercept are not entirely appropriate either. The reason stems from the fact that what should be tested here is not separate tests of $\alpha = 0$ and $\beta = 1$, rather simultaneous testing of these hypothesis is required, i.e.,

$$\text{(6-3)} \qquad \left\{ \begin{array}{l} H_0 : \ \alpha = 0 \wedge \beta = 1 \\ H_a : \ at\ least\ one\ isn't \end{array} \right\}$$

To alleviate these problems, conducting a joint F-test is suggested, e.g., Harrigan, et al. (1980). In the current project, a restricted joint F-test was conducted. Details of this test is explained in chapter 5, and **(6-3)** above represents the relevant null and alternative hypothesis. The appropriate F-values are calculated from the regression data via utilization of equation **(5-6)**, and are presented in table 6-6 below.

Using 1% level of significance in the case of direct coefficients, one will fail to reject the null hypothesis only in the case of RAS, RECRAS, and RASLAG. In all other cases the null hypothesis can be rejected. In other words the intercept and the slope are not simultaneously equal to

Table 6-6

Calculated F-statistic

Direct and Inverse Coefficients

Estimation Technique	F-Statistic Direct	F-Statistic Inverse
NAIVE	16.16	115.79
RAS	0.41	2.12
RECRAS	2.13	15.74
PROPVA	654.71	309.69
FRIED	5.08	6.33
RECLAG	113.51	181.30
RASLAG	0.41	2.12
RERALA	5.53	1.13
ALMON	41.01	414.45

zero and one, indicating that the compared matrices are not statistically similar. For the inverse matrix the null hypothesis can not be rejected for RAS, RASLAG, and RERALA only. Hence, suggesting similarity of estimated matrices and the actual one only in these cases.

If one chooses the level of significance at 0.5%, then one will fail to reject the null hypothesis in case of RAS, RECRAS, RASLAG, and FRIED. In the case of RERALA, the null hypothesis is barely rejected. But in all other cases one still rejects the null hypothesis, with the consequences being the same as before. Namely, absence of similarity between actual matrices and the updated versions. In the case of inverse matrices the good estimates are obtained via RAS, RASLAG, and RERALA. In all other cases the null hypothesis is rejected with the same implications as before.

It should be reiterated again that this test is really a measure of general tendencies and not strict cell by cell comparison. As such, it does not take into account the distance between individual cell. Thus, it may give an erroneous impression of a poor (good) estimate while in actuality the estimate might be a satisfactory (poor) one. This is one reason why other tests of closeness are included in evaluation of the performance of various techniques.

It also worth remembering that quite possibly in some of these regressions a block of coefficients are estimated accurately, while the overall estimates are not. The probe of this possibility is deferred to chapters 7 and 8.

From the evidence presented so far, one might be tempted to conclude that accurate estimates of matrices are obtained through utilization of some of the techniques (i.e., $\alpha = 0$ and $\beta = 1$). This conclusion, however, may be premature since the errors may be in such a way that neutralize one another and present an erroneous impression of high accuracy. To investigate this possibility, several other tests are needed. These tests and their respective results are presented below.

The ranking of the techniques based on the joint F-test for direct coefficients is RAS & RASLAG, RECRAS, FRIED, RERALA, NAIVE, ALMON, RECLAG, and PROPVA. For the inverse coefficients the ranking would be RERALA, RAS & RASLAG, FRIED, RECRAS, RECLAG, NAIVE, PROPVA, and ALMON.

6.2.5) **Chi Square Test**. As a supplementary step in testing the closeness of updated matrices and the benchmark table the chi-square goodness of fit test is employed. This test was conducted at two levels. First the test, for each estimate, was performed over the entire matrix (i.e., only one class). The next level was to divide each matrix into forty classes and conduct the test accordingly. The distribution of actual and estimated coefficients are, for comparative purposes, presented in tables C-1 to C-20 and figures F-1 to F-20 of Appendices (C) and (F).

In each case the null hypothesis of similarity of distributions of the estimated and actual coefficients in each class was tested against the alternative hypothesis that they do not follow the same distribution. The χ^2 statistic at 1% and 0.5% levels of significance is utilized for both series of classifications. Then, if the calculated χ^{*2} is less than the critical value of χ^2 for a prescribed level of confidence, the null hypothesis may not be rejected, otherwise the null hypothesis can be rejected. If one fails to reject the null hypothesis, then it may be inferred that the estimated and actual coefficients follow the same distribution, which in turn could be an indication of a good estimate. It is clear that the smaller the χ^2 the better is the fit.

Under forty classes scheme (39 degrees of freedom), the relevant critical values of χ^2 for 1% and 0.5% are 62.428 and 65.476 respectively. The same critical values for one class scheme (5040 degrees of freedom) are 5301.84 and 5325.56 respectively. Calculated χ^2 values for various updating methods are presented in tables 6-7 and 6-8. Comparison of these values with the critical values will allow one to make some inferences about the distribution of the estimated coefficients. As it can be seen, in the case of forty classes and direct coefficients, for both 1% and 0.5% levels, one will reject the null hypothesis for all methods except FRIED, indicating absence of similarity of distribution between the estimates and actual data.

In the same case and for inverse coefficients, however, at both 1% an 0.5% one will fail to reject the null hypothesis in all but the ALMON case. This fact, then indicates similarity of distribution among the estimated and actual inverse coefficients.

The ranking of the techniques based on forty class χ^2

TABLE 6-7

CHI SQUARE TEST, FORTY CLASSES

	DIRECT COEFFICIENTS	INVERSE COEFFICIENTS
NAIVE	315.53724	42.939330
RAS	326.60316	22.508746
RECRAS	301.82449	21.173312
PROPVA	313.74114	49.296278
FRIED	59.067852	17.979173
RECLAG	331.69679	42.638662
RASLAG	326.60316	22.508746
RERALA	306.37796	25.968099
ALMON	305.42908	367.91917

is FRIED, RECRAS, ALMON, RERALA, PROPVA, NAIVE, RAS & RASLAG, and RECLAG. The same for the inverse coefficients is FRIED,RECRAS, RAS & RASLAG, RERALA, RECLAG, NAIVE, PROPVA, and ALMON.

In the case of individual coefficients, the calculated values, (χ^{*2}'s), for both direct and inverse coefficients are far less than the critical values at 1% and 0.5%. This will lead one to fail to reject the null hypothesis for direct and inverse coefficients under all methods, which

TABLE 6-8

CHI SQUARE TEST, INDIVIDUAL COEFFICIENTS

	DIRECT COEFFICIENTS	INVERSE COEFFICIENTS
NAIVE	78.24639	20.45808
RAS	54.33973	11.46280
RECRAS	56.12835	15.93332
PROPVA	79.38877	32.84723
FRIED	60.05370	11.63787
RECLAG	130.80943	32.81232
RASLAG	54.33974	11.46280
RERALA	55.98847	14.16662
ALMON	12.54586	32.03400

indicates strong similarity among the distributions of estimated and actual coefficients.

The ranking for direct coefficients based on one class χ^2 is ALMON, RAS & RASLAG, RERALA, RECRAS, FRIED, NAIVE, PROPVA, and RECLAG. The same ranking for inverse coefficients is RAS & RASLAG, FRIED, RERALA, RECRAS, NAIVE, ALMON, RECLAG, and PROPVA.

Thus, as it is evident, the results are somewhat ambiguous. This stems from the nature of this test as well as the disposition of estimation techniques. A brief explanation of this point is in order.

First, as was mentioned before, matrix comparison can be conducted at partitive as well as holistic levels. In the first case, one is concerned with cell by cell accuracy of the estimated matrices, while in the second case overall accuracy of each estimate is of concern. It was also noted that, due to the structure of the minimands of the techniques used in this study, a *priori* expectation is that all of these techniques provide more accurate estimates as one moves from partitive comparison toward holistic one. In fact this was the reason for omitting the total output and multiplier comparisons in this study, since good estimates were guaranteed in these cases, hence rendering evaluation of techniques meaningless.

Having these in mind, then it is not surprising to see that one fails to reject the null hypothesis in the case of inverse coefficients (i.e., a good fit between the actual inverse coefficients and the estimated ones.) while rejecting the null hypothesis in the case of direct coefficients. Since, as mentioned earlier, comparison of direct coefficients is the most partitive level of comparison of two matrices while comparison of inverse coefficients is a step closer to holistic comparison, thus more likely to yield closer results to the benchmark table.

Second, the chi-square test is extremely sensitive to the classification. It is possible to manipulate the class width to arrive at the desired χ^2. This is evident from comparison of tables 6-7 and 6-8. In the first case, the estimated coefficients were divided into forty classes of equal width, and result was rejection of null hypothesis in the case of direct coefficients. However, in the second case when all coefficients were grouped in one class, the null hypothesis could no longer be rejected. This indicates an overall goodness of fit but not an accurate fit over specific size distributions. Therefore, the possibility of manipulating the results via alteration of class width or classification of coefficients in classes that do not have equal width, drastically reduces the power of this test.

Third, in any I-O table there are cells with zero values. Now, if for a given coefficient the initial value is zero and the estimated value is non-zero, to avoid division by zero, most researchers discard these pairs in the process of calculating the χ^2 statistic (e.g., Morrison and Smith, 1974; Harrigan et al, 1980; Butterfield and Mules, 1980, and the current study). In other words, in the calculation of χ^2 zeros will replace what is really not defined, hence drastically reducing the value of the statistic and possibly erroneous impression of a close estimate. The same problem can arise if an estimate generates negative numbers. For these negatives are replaced by zeros, which in turn give rise to the problem just mentioned. This problem is evident in the case of ALMON's direct coefficient. The ALMON method is judged by all other tests to provide a poor estimate of actual matrix. However, the chi-square statistic associated with ALMON model is the smallest χ^2 in the table 6-8, giving an indication of a close estimate or at least an estimate that is superior to those of other methods. An examination of table 6-1, however, reveals that this technique generated a large number of negative coefficients. Since these negative numbers were replaced by zeros, the low value of χ^2 was obtained, leading to the wrong impression that ALMON model outperformed other techniques.

To overcome this last problem, one can replace the zeros with a very small values. But this remedy may be worse than the disease, for chi-square statistic is very sensitive to small numbers in denominator and this replacement will increase its value substantially and will bias the results. A better solution is to make the comparison over a size distribution of the coefficients. This, however, will bring back the problem of classification and the class width alluded to above.

Forth, the chi-square statistic can not capture the structural differences between matrices and solely rely on the numerical values, i.e., the value of coefficients in the matrix and not their location. This, as noted by Israilevich (1986), can cause that two estimated matrices to have identical χ^2 values, hence judged equally good in estimation power, while their structure is extremely different. Which means one of them may resemble the benchmark table quite well while the other drastically differs from the base matrix, but they were both ranked equally. By the same token, it is also conceivable to have two updated matrices where one has smaller χ^2 value hence judged a better estimate, while in reality the estimate with higher χ^2 value resembles the actual table more closely.

Finally, if in the estimation process some coefficients are relocated from one class to another, the χ^2 value may change substantially and misleading results may be obtained. This is very possibly a source contributing to the rejection of null hypothesis in the forty class scheme under the current study.

The realization of these shortcomings led to inclusion of other tests in this project and usage of chi-square statistic test only as a relative measure of distance instead of an absolute goodness of fit test.

6.2.6) **Mean Absolute Deviation**. This statistic is used as summary measure of distance between actual and estimated coefficients. The smaller the value of MAD, the better is the estimate, and zero value of MAD is associated with a perfect fit. MAD is heavily influenced by presence of large errors, hence it should be considered along side the mean and standard deviation of estimates.

The first columns of tables 6-9 and 6-10 present the MAD statistic for the estimation techniques utilized here. These statistics range from low of 0.0019 for RAS to high of 0.02609 for ALMON. They are all seem to be small, but these values in themselves can not be judged, for there is no objective threshold for this purpose. They may be used only for measuring relative performances of different techniques. Accordingly, the ranking of the models for direct coefficients is RAS & RASLAG,FRIED, RERALA, RECRAS, NAIVE, PROPVA, RECLAG, and ALMON. The same ranking for inverse coefficients is RAS & RASLAG, FRIED, RERALA, RECRAS, NAIVE, RECLAG, PROPVA, and ALMON.

6.2.7) Standardized Total Percentage Error. This is another measure of distance between estimated and actual coefficients. It shows the average of absolute differences among estimated and actual tables as percentage of the mean of the actual table. It is very sensitive to zeros and small values. A perfect fit will make this statistic to assume value of zero.

Columns two of tables 6-9 and 6-10 show this statistic for direct and inverse coefficients respectively. All methods, excluding ALMON, seem to provide reasonable values of STPE. Based on these tables the STPE for direct coefficients range from low of 0.26347 for RAS to high of 2.74233 associated with ALMON, yielding the ranking of the

TABLE 6-9

ABSOLUTE AND RELATIVE MEASURES OF FORECASTING ACCURACY

ACTUAL AND PREDICTED DIRECT COEFFICIENTS

	MAD	STPE	RMS	U	UM	US	UC
NAIVE	0.00222	0.29379	0.07492	0.12456	0.00002	0.00005	0.99993
RAS	0.00199	0.26347	0.05899	0.09800	0.00000	0.00008	0.99992
RECRAS	0.00218	0.28894	0.07303	0.12046	0.00000	0.00022	0.99977
PROPVA	0.00235	0.31140	0.08623	0.15137	0.00009	0.00139	0.99852
FRIED	0.00205	0.27168	0.06150	0.10202	0.00000	0.00011	0.99989
RECLAG	0.00259	0.34357	0.09030	0.14448	0.00010	0.00149	0.99841
RASLAG	0.00199	0.26347	0.05899	0.09800	0.00000	0.00008	0.99992
RERALA	0.00213	0.28226	0.07128	0.11713	0.00000	0.00036	0.99964
ALMON	0.02069	2.74233	0.63007	0.65556	0.00001	0.00482	0.99518

estimation methods to be RAS & RASLAG, FRIED, RERALA, RECRAS NAIVE, PROPVA, RECLAG, and ALMON. Same data set for inverse coefficients indicates the value of STPE to range from low of 0.10664 for RAS to high of 0.52909 for ALMON with the relevant ranking of RAS & RASLAG, FRIED, RERALA, RECRAS, NAIVE, RECLAG, PROPVA, and ALMON.

6.2.8) Root Mean Square. Root mean square of the elements of the matrix of difference between the estimated and actual matrices, dubbed as RMS, can serve as a measure

TABLE 6-10

ABSOLUTE AND RELATIVE MEASURES OF FORECASTING ACCURACY

ACTUAL AND PREDICTED INVERSE COEFFICIENTS

	MAD	STPE	RMS	U	UM	US	UC
NAIVE	0.00435	0.14377	0.12210	0.04976	0.00015	0.00025	0.99960
RAS	0.00323	0.10664	0.08773	0.03547	0.00000	0.00000	1.00000
RECRAS	0.00413	0.13646	0.13019	0.05272	0.00002	0.00000	0.99998
PROPVA	0.00541	0.17887	0.16551	0.06809	0.00037	0.00071	0.99891
FRIED	0.00330	0.10908	0.09120	0.03690	0.00000	0.00000	1.00000
RECLAG	0.00531	0.17531	0.17242	0.06835	0.00008	0.00137	0.99855
RASLAG	0.00323	0.10664	0.08773	0.03547	0.00000	0.00000	1.00000
RERALA	0.00374	0.12346	0.10944	0.04413	0.00000	0.00008	0.99992
ALMON	0.01601	0.52909	0.41030	0.17208	0.00021	0.00053	0.99925

through which one can determine how closely the individual estimated coefficients resemble their actual counterparts. In ranking of the updating techniques, a lower value of RMS indicates a better performance for the method. A zero value of RMS signifies a perfect estimate.

Third columns of tables 6-9 and 6-10 present this statistic for direct and inverse coefficients estimated via various methods. Here again, baring ALMON, all methods produce small RMS values. In the case of direct coefficients, the value of RMS ranges from low of 0.05899 associated with RAS to high of 0.63007 for ALMON method. The ranking, accordingly, will be RAS & RASLAG, FRIED, RERALA, RECRAS, NAIVE, PROPVA, RECLAG, and ALMON. For inverse coefficients, the RMS assumes the values from 0.08773 for RAS to 0.4103 for ALMON, thus yielding ranking of the techniques as RAS & RASLAG, FRIED, RERALA, NAIVE, RECRAS, PROPVA, RECLAG, and ALMON.

6.2.9) Theil's U. The U statistic is constructed via utilization of RMS, and is used for comparison of estimated and actual coefficients. Its value fluctuates between one and zero, with zero representing a perfect estimate.

Columns four of the tables 6-9 and 6-10 represent the U statistic for direct and inverse coefficients of various estimation methods, and indicate acceptable values for U. The U in the case of direct coefficients ranges from low of 0.09800 for RAS to high of 0.65556 for ALMON. With the exception of direct ALMON, none of the methods seem to have a high U. Accordingly, the ranking of the techniques will be RAS & RASLAG, FRIED, RERALA, RECRAS, NAIVE, RECLAG, PROPVA, and ALMON. For inverse coefficients the low and high are 0.03547 and 0.17208 for RAS and ALMON respectively, yielding the ranking of RAS & RASLAG, FRIED, RERALA, NAIVE, RECRAS, PROPVA, RECLAG, and ALMON.

6.2.10) UM. The U statistic can be decomposed into three component parts, each shedding some lights on the estimated tables and their corresponding method of estimation. UM is one of the three components and represent the extent of systematic error and measures the extent of deviation of the average values of estimated and actual coefficients. The value of UM can vary between zero and one, with zero implying absence of systematic error in the estimate.

The UM values for estimation techniques utilized here are tabulated in the fifth columns of tables 6-9 and 6-10, and they all seem to be reasonably low. For direct coefficients this value ranges from zero to 0.00009 providing the ranking of RAS & RASLAG & RECRAS & FRIED & RERALA, ALMON, NAIVE, PROPVA, and RECLAG.

In the case of inverse coefficients the UM values oscillate from zero to 0.00037 with the corresponding ranking of RAS & RASLAG & FRIED & RERALA, RECRAS, RECLAG, NAIVE, ALMON, and PROPVA.

6.2.11) US. The second component of the U statistic, US, is the variance proportion of U and represents the technique's ability to generate an estimate that emulates the variation in the actual table. Its presence indicates inefficiency in measurement of degree of association between actual and estimated data, i.e., either the actual coefficients have experienced a fluctuation that the estimated set fails to reflect, or a fluctuation is reflected in the estimates that really did not occur in the actual data. The value of US can range between zero and one, with zero being most efficient case.

The sixth columns of tables 6-9 and 6-10 provide this statistic for the estimation techniques used in this study. These values appear to be low and desirable. The value of US for direct coefficients fluctuates from 0.00005 to 0.00482 leading the models to be ranked as NAIVE, RAS & RASLAG, FRIED, RECRAS, RERALA, PROPVA, RECLAG, and ALMON. In the case of inverse coefficients values of US ranges from zero to 0.00137 giving the ranking of RAS & RASLAG & RECRAS & FRIED, RERALA, NAIVE, PROPVA, ALMON, and RECLAG.

6.2.12) UC. The third proportion of Theil's U is the covariance proportion, UC, which measure the unsystematic error. It is the residual of U after UM and US are accounted for. Since UM + US + UC = 1, and the best values of UM and UC are zeros, the optimal value of UC is one.

The seventh columns of tables 6-9 and 6-10 provide the values of UC for the methods used in this study, and all values seem to be high and acceptable. In the case of direct coefficients UC varies from low of 0.99518 to high of 0.99993, making the ranking to be NAIVE, RAS & RASLAG, FRIED, RECRAS, RERALA, PROPVA, RECLAG, and ALMON. Inverse coefficients have the UC ranging from 0.99855 to one with ranking of RAS & RASLAG & FRIED, RECRAS, RERALA, NAIVE, ALMON, PROPVA, and RECLAG.

It is evident from tables 6-9 and 6-10 that all models for both direct and inverse coefficients have very good, and in some cases the ideal, values of UM, US, and UC. It must be emphasized again that this fact per se does not make these estimation methods "good techniques." Because these statistics do not have any meaning in isolation and their values must be evaluated in conjunction with the value of Theil's U. These statistics merely decompose the Theil's U, and within that context alone provide some insight into the variation due to bias of estimation as well as the variance and covariance of the U statistics.

6.2.13) **Mean, Standard Deviation, and the Maximum Value**. As additional supplementary measures, in the present study these statistics are utilized in comparison of estimated and actual matrices. A good estimate should have a mean and a standard deviation close to those of the actual table. Comparison of the maximum values of estimated and actual coefficients should provide a rough measure of performance of an estimation procedure.

Tables 6-11 and 6-12 tabulate these statistics for various updating techniques alongside with their counterparts from the actual benchmark tables for direct and inverse coefficients.

In the case of inverse coefficients the situation is

TABLE 6-11

MEAN AND STANDARD DEVIATION OF PREDICTED DIRECT COEFFICIENTS

METHOD	MEAN	STANDARD DEVIATION	MAXIMUM
ACTUAL	0.00754	0.03465	0.70512
NAIVE	0.00724	0.03519	0.90439
RAS	0.00754	0.03518	0.71255
RECRAS	0.00739	0.03574	0.83913
PROPVA	0.00672	0.03144	0.63717
FRIED	0.00754	0.03529	0.71581
RECLAG	0.00664	0.03814	0.76104
RASLAG	0.00754	0.03518	0.71255
RERALA	0.00744	0.03601	0.76860
ALMON	0.00594	0.07839	1.84481

similar. RAS, RASLAG, and FRIED have means that exactly match the mean of actual table. RECRAS, RERALA, RECLAG, NAIVE, PROPVA, and ALMON respectively underestimate the mean of inverse coefficients by 2%, 4%, 5%, 5%, 11%, and 20%.

The standard deviations of, RAS, RASLAG, FRIED,and RECRAS, are identical to that of the actual inverse matrix and RERALA almost matches the same value. The same statistic

TABLE 6-12

MEAN AND STANDARD DEVIATION OF PREDICTED INVERSE COEFFICIENTS

	MEAN	STANDARD DEVIATION	MAXIMUM
ACTUAL	0.03026	0.14357	2.05370
NAIVE	0.02879	0.14163	1.84332
RAS	0.03026	0.14370	2.04409
RECRAS	0.02974	0.14331	1.89921
PROPVA	0.02706	0.13915	2.15907
FRIED	0.03026	0.14349	2.05249
RECLAG	0.02876	0.14995	2.11685
RASLAG	0.03026	0.14370	2.04409
RERALA	0.03013	0.14452	2.05111
ALMON	0.02429	0.13409	1.44398

for NAIVE, PROPVA, RECLAG, and ALMON deviate from the actual one by 1%, 3%, 4%, and 7% respectively.

With regards to the maximum values, RAS, RASLAG, FRIED, and RERALA have the same maximum value as the actual matrix. RECLAG, PROPVA, RECRAS, NAIVE, and ALMON deviate from this value by 3%, 5%, 8%, 10%, and 30% respectively.

Thus, according to the mean, standard deviation, and maximum value of the coefficients, it seems that some of the methods have the same or very close central tendency and dispersion as the actual data, which indicate close fits.

6.3) SUMMARY OF THE RESULTS:

Tabular summary of the rankings of the estimation methods as gauged by the specified criterions for both direct and inverse coefficients is presented in tables 6-13 and 6-14 below. It should be mentioned that not all the criterions *per se* will lead to ranking of the estimation methods (e.g., Wilcoxon, standard deviation, etc.). Thus, in these cases, some rather arbitrary rule (such as narrowing the acceptance intervals) was employed in order to allow simultaneous presentation of most of pertinent information for comparison and *ranking* of the estimation techniques.

Numbers at the cross section of any given method and a given criterion refers to the ranking of the former via utilization of the latter. In cases that ranks were tied, same number appears in front of all methods of equal rank, and the next method in ranking was assigned its true position, i.e., 1, 1, 1, 4, 5, indicates the first three techniques have equal ranking and the next best technique assumes the rank of four.

Explanation of symbols and acronyms follows:

N.N = number of estimated negative coefficients.

W.U = Wilcoxon's U for rank sum test.

$\theta/5$ = coefficient of equality at 5% level.

$\theta/20$ = coefficient of equality at 20% level.

μ_θ = mean of coefficient of equality.

σ_θ = standard deviation of coefficient of equality.

MAX_θ = maximum value of coefficient of equality.

F = F-value, overall test of significance.

t_α = t-value, test for intercept value.

t_β = t-value, test of slope coefficient.

R^2 = regression line's coefficient of determination.

$\chi^{2/1}$ = chi-square, one class.

$\chi^{2/40}$ = chi-square, forty classes.

J.F = F-value for the joint test.

MAD = mean absolute deviation of estimated coefficients.

STP = standardized total percentage errors of estimates.

RMS = root mean square of estimates.

U = Theil's U statistic.

UM = mean of Theil's U.

US = standard deviation of Theil's U.

UC = covariance of Theil's U.

μ_a = mean of estimated coefficients.

MAX_a = maximum value of estimated coefficients.

σ_a = standard deviation of estimated coefficients.

TABLE 6-13

SUMMARY RANKINGS OF ESTIMATION METHODS

DIRECT COEFFICIENTS

METHOD	N.N	W.U	θ/5	θ/20	μ_θ	σ_θ	MAX_θ	F
NAIVE	1	3	7	4	8	8	8	6
RAS	1	1	1	1	5	5	4	1
RECRAS	5	5	3	5	3	3	3	5
PROPVA	1	5	6	7	7	7	7	8
FRIED	5	3	5	3	4	4	6	3
RECLAG	5	8	8	8	1	1	1	7
RASLAG	1	1	1	1	5	5	4	1
RERALA	5	7	4	6	2	2	2	4
ALMON	9	9	9	9	9	9	9	9

TABLE 6-13 CONTINUED

SUMMARY RANKINGS OF ESTIMATION METHODS

DIRECT COEFFICIENT

METHOD	t_α	t_β	R^2	$\chi^{2/1}$	$\chi^{2/40}$	J.F	MAD	STP
NAIVE	7	6	6	7	6	6	6	6
RAS	2	3	1	2	7	1	1	1
RECRAS	6	1	5	5	2	3	5	5
PROPVA	5	9	8	8	5	9	7	7
FRIED	1	2	3	6	1	4	3	3
RECLAG	9	8	7	9	9	8	8	8
RASLAG	2	3	1	2	7	1	1	1
RERALA	8	5	4	4	4	5	4	4
ALMON	4	7	9	1	3	7	9	9

TABLE 6-13 CONTINUED

SUMMARY RANKINGS OF ESTIMATION METHODS

DIRECT COEFFICIENT

METHOD	RMS	U	UM	US	UC	μ_a	MAX_a	σ_a
NAIVE	6	6	7	1	1	6	8	1
RAS	1	1	1	2	2	1	1	1
RECRAS	5	5	1	5	5	5	7	5
PROPVA	7	8	8	7	7	8	6	7
FRIED	3	3	1	4	4	1	3	1
RECLAG	8	7	9	8	8	7	4	8
RASLAG	1	1	1	2	2	1	1	1
RERALA	4	4	1	6	6	4	5	6
ALMON	9	9	6	9	9	9	9	9

TABLE 6-14

SUMMARY RANKINGS OF ESTIMATION METHODS

INVERSE COEFFICIENTS

METHOD	N.N	W.U	θ/5	θ/20	μ_θ	σ_θ	MAX_θ	F
NAIVE	1	6	6	6	1	8	8	5
RAS	1	1	1	1	4	2	4	1
RECRAS	1	5	5	5	2	5	2	6
PROPVA	1	7	7	7	7	7	7	8
FRIED	1	1	3	3	6	6	6	3
RECLAG	1	8	8	8	8	1	3	7
RASLAG	1	1	1	1	4	2	4	1
RERALA	1	4	4	4	3	4	1	4
ALMON	9	9	9	9	9	9	9	9

TABLE 6-14 CONTINUED

SUMMARY RANKINGS OF ESTIMATION METHODS

INVERSE COEFFICIENTS

METHOD	t_α	t_β	R^2	$\chi^{2/1}$	$\chi^{2/40}$	J.F	MAD	STP
NAIVE	7	6	5	6	7	7	6	6
RAS	1	1	1	1	3	2	1	1
RECRAS	5	5	6	5	2	5	5	6
PROPVA	8	8	8	9	8	8	8	8
FRIED	3	4	3	3	1	4	3	3
RECLAG	9	7	7	8	6	6	7	7
RASLAG	1	1	1	1	3	2	1	1
RERALA	4	3	4	4	5	1	4	4
ALMON	6	9	9	7	9	9	9	9

TABLE 6-14 CONTINUED

SUMMARY RANKINGS OF ESTIMATION METHODS

INVERSE COEFFICIENTS

METHOD	RMS	U	UM	US	UC	μ_a	MAX_a	σ_a
NAIVE	5	5	7	6	6	7	8	6
RAS	1	1	1	1	1	1	1	2
RECRAS	6	6	5	1	4	4	7	4
PROPVA	7	7	9	7	8	8	6	7
FRIED	3	3	1	1	1	1	1	1
RECLAG	8	8	6	9	9	6	5	8
RASLAG	1	1	1	1	1	1	1	2
RERALA	4	4	1	5	5	5	1	5
ALMON	9	9	8	8	7	9	9	9

In interpreting tables 6-13 and 6-14, a great deal of caution must be exercised.

First, the place of a given method in ranking by a closeness test in and of itself is not sufficient in rendering a judgment on that technique, and the ranking based on different tests do not have equal importance. Poor performance via some tests can be tolerated while failure in other tests may be detrimental to a given technique. For instance, if a technique, while passing the acceptable threshold, is ranked low by the F-test but generates no negative values for coefficients should be judged superior to a test with very high F-value and some negative coefficients.

Second the low or high ranking of some methods are explained rather easily and does not necessarily reflect the method's low or high estimation power. That is to say, a particular method might be ranked high or low not due to estimation power of the technique, but due to other circumstances such as the make up the original table or some shortcomings of statistical tests of closeness, etc.

Third, some methods will have very accurate estimates for a set of coefficients and rather large deviation on others, while different methods may have no set of estimation with high degree of accuracy but enjoy modest deviation throughout the estimated matrix. This phenomenon may drastically bias the tests and their results.

These and similar points must be kept in mind before one can proceed to evaluate and rank the estimation methods.

Overall, considering all tests and criterions utilized, it appears that some of the estimation methods replicate the actual data very closely while others do not fare as well. Almon method can be eliminated immediately by virtue of its poor showing in all tests, specially due to very large number of negative coefficients that it generates. The only exception to ALMON's poor ranking is the χ^2 tests. This good result is due to nothing but those negative coefficients that create the illusion of a good estimate. The non-negativity and Wilcoxon tests seriously damage the RECLAG and PROPVA models, the end for these models comes in the hand of regression analysis, when one tests for values of intercept and slope coefficients. The meager display of these methods in other tests warrants this conclusion. The good results of RECLAG in the mean and standard deviation of the coefficient of equality merely reflects a good distribution of bad results, i.e., nullification of most of the errors by positive and negative deviations. It should be emphasized again that no method is "good" or "bad" in absolute term. Once a technique passed a certain set of criterion, the "goodness" determination depends on the circumstances governing any given experiment, i.e., the resources available, the specific objectives of the experiment, and the degree of tolerance for error. It is the purpose of this study to *rank* the estimation techniques, and it is on this basis alone that a given technique is evaluated. RERALA and NAIVE methods, although pass almost all the required tests, in relative terms fall short of exhibiting good results. The exceptions, again, being due to cancellation of positive and negative errors rather than the power of these methods. RECRAS method also seems to yield reasonable general results with some inadequacy in particular areas. FRIED method displays good results, albeit with some explainable shortcomings, and can be ranked as the third method among the techniques utilized in this experiment. RAS and RASLAG appear to be the best of all techniques used here. These methods, however, do not score very high based on forty class chi-square and coefficient of equality (and its derivatives) criterions. This is probably due to the sensitivity of chi-square test to classification, existence of very small coefficients in the actual tables, as well as the way RAS procedure works. When the original coefficient is very small, even a rather negligible adjustment via RAS (or RASLAG) procedure amounts to a high percentage change and deteriorates RAS' (or RASLAG's) ranking.

Therefore, as noted above, it seems that some of the methods employed here produce rather accurate results. prudence, however, must be the guide here and hasty conclusions must be avoided.

To begin with, the time span between the two tables used for the current experiment is not very long and the period 1966-1972 in the Soviet Union was not marked by rapid technological changes, which might partly explain rather close estimates.

Furthermore, the results are not always conclusive and there is a certain degree of ambiguity, and at times even contradictions, in conclusions obtained via these tests. One should constantly be reminded that the indication of closeness given by these results is really more on the general level than exact cell by cell basis, and these tests should really be used for *ranking* of the updating methods rather than an absolute measure of *accuracy* of estimated tables.

Moreover, there are number of cells with a very small coefficients, and even a small error on these coefficients amount to a high magnitude in relative terms, hence indicating scanty estimation when the results are actually reasonable. Additionally, the positive and negative deviations tend to balance each other out, hence escaping detection by the closeness tests and indicating good estimates despite an actual poor performance. It is quite possible, and in fact likely that many small coefficients were not accurately estimated, but this was covered by relative accuracy of medium to large coefficients.

The bias introduced in the results obtained via different estimation methods and tests of closeness must not be ignored. It is likely that a method estimating a good number of cells with high degree of accuracy be ranked low due to its sharp deviation in estimating a subset of coefficients, or conversely a technique may be ranked high due to rather moderate errors throughout the table. These and similar possibilities along with the structural biases inherent in some tests of closeness can skew the results with some quite misleading and serious consequences.

None of these points, however, diminishes the usefulness of either estimation techniques or the statistical tests. These techniques may be very helpful and sometimes the only alternative available to given researchers. The points raised here are precautionary in nature and mentioned merely as an attempt to put the matter in proper perspective. They are not, nor intended, as whole sale dismissal of neither non-survey estimation and updating techniques nor tests of closeness for matrices.

CHAPTER SEVEN

INCORPORATION OF EXOGENOUS INFORMATION

"I don't know why. It seemed a good thing to be doing. It seemed to have some meaning."

John Steinbeck

7.1) INTRODUCTION:

During updating and projecting process of I-O tables quite often researchers find additional data at their disposal over and above what is needed by the updating techniques. The exogenous information, is usually result of an independent survey or forecast of a given segment of the target year, conducted for other purposes and available to I-O researchers. Moreover, in some instances factors such as availability of resources (time, budget, staff, etc.), importance of a given sector, sudden changes in technology, or sensitivity of the overall forecast to a given segment of the economy warrants exogenous estimation or calculation of some parts of the I-O table. The question, then, is what to do with the exogenous information and how they should be treated in construction of the I-O table.

It is believed by many analysts such as Lecomber (1964), Allen (1974) and (1975), Allen and Lecomber (1975), Jensen and West (1980), and Miller and Blair (1985), etc.; that under these circumstances, it may be advisable, and indeed sometimes necessary, to utilize such additional information in the updating process. The inclusion is expected to improve the estimates, and in fact many researchers report improvement in the results due to employment of exogenous data. However, as noted by authors like Miernyk (1976), and Israilevich (1986), the incorporation of additional data sometimes can exacerbate the results. That is, utilization of the exogenous information could also lead to a less accurate updated matrix. While Miernyk used this possibility as a basis of his discounting of mechanical updating procedures, Israilevich attempted an explanation of the phenomena and conditions that lead to such results, as well as proposing some remedies for the problem.

It should also be noted that most studies take the exogenous estimates to be reliable. Lecomber (1964), however, questions this assumption and suggests an algorithm that can take into account varying degrees of reliability of information. His scheme was tested by Allen and Lecomber (1975), with marked improvement in the results.

7.2) MODIFIED UPDATING METHODS:

In the present project, to address the issue of incorporation of additional information, exogenous data is utilized to modify three updating procedures, namely NAIVE, RAS, and Lagrangian, under three different schemes. These methods are selected both in the interest of continuity and compatibility, as well as due to their relative performance accuracy as indicated in chapter six. In addition, one variant of a rather mechanical modification scheme, namely Residual Minimum method is utilized, for its relative ease and compatibility with other methods. A brief description of these procedures follows.

7.2.1) **Modified RAS Procedure**. The Simple RAS method cannot incorporate any data about the target year matrix except the marginal totals. However, many algorithms have been suggested to modify the RAS method and allow the usage of additional data. Among these suggestions Paelink and Waelbroeck (1963, cited in Lecomber, 1975), Omar (1967), Grandville et al. (1968), Bacharach (1970), Lecomber (1964) and (1975), Allen (1974), Allen and Lecomber (1975), Hewings and Janson (1980), Jensen (1980), Jensen and West (1980), Hewings and Syverson (1982, cited in Polenske et al. 1986), Alexeev (1983), and Israilevich (1986), may be mentioned.

Specifically, availability of exogenous information boils down to knowledge of z_{ij}^{t} for a particular cell.

Since X_j^t is known then, by virtue of the basic I-O relationship, a_{ij}^{t} is also known. Incorporation of these known coefficients into the original RAS method constitutes the heart of the Modified RAS technique.

In doing so, one must subtract the known z_{ij}^{t} 's from the corresponding U_i^t and V_j^t to obtain new relevant marginals. The base year's technology matrix A^0 also must be modified by replacing zeros in place of all known coefficients. The regular RAS procedure, then, can be applied to the resultant matrix ($\overline{A}^0$) and new row and column marginals ($\overline{U}_i^t$ and $\overline{V}_j^t$) to obtain estimates of the remaining non-zero coefficients. The next step is to insert the exogenously determined coefficients in the updated matrix in order to arrive at the complete updated I-O coefficients for the target year. The zero preservation property of the RAS technique will guarantee that the cells corresponding to the exogenously determined coefficients remain zero in the process.

Formally, the procedure can be expressed as:

(7-1) $$A^t = [C] + [\hat{R}]\ [A^0 - C]\ [\hat{S}]$$

where [C] is an (n) by (n) null matrix with c_{ij}^t 's replaced by the exogenously determined coefficients of the target year. Denoting the middle term of the second argument of this equation by $\overline{A}^0$ one will get:

(7-2) $$A^t = [C] + [\hat{R}]\ [\overline{A}^0]\ [\hat{S}]$$

Then, the standard RAS procedure can be applied to the modified matrix $\overline{A}^0$ subject to:

$$\text{(7-3)} \qquad \sum_{j=1}^{n} r_i \, x_{ij}^{0} \, s_j = u_i^t - \sum_{j=1}^{n} (c_{ij} \, X_j)$$

and:

$$\text{(7-4)} \qquad \sum_{i=1}^{n} r_i \, x_{ij}^{0} \, s_j = v_j^t - \sum_{i=1}^{n} (c_{ij} \, X_i)$$

to get the update of the modified base matrix. Summation of this updated matrix and the C matrix which was constructed based on exogenous information will yield the final estimate of the target year's matrix.

A priori, the expectation is that the inclusion of additional data should improve the results. However, the possibility of obtaining inferior results can not be ruled out. This paradoxical possibility stems from the very nature of the RAS procedure and the modification process. For incorporation of additional data in the original matrix amounts to insertion of zeros in their place in the modified base year matrix, that is to say, the modified matrix contains less non-zero cells than the original matrix. When the RAS method is applied to it, a smaller number of cells must be adjusted in such a way to satisfy the required marginal totals. Hence, the possibility arises of higher degree of individual cells deviation from their true values. This in turn, might lead to inferior measure of overall accuracy as compared to the case of no exogenous data.

Israilevich (1986) explains the same phenomena by demonstrating that the imposition of a constraint on a cell or a set of cells will introduce a change throughout the whole matrix. In other words, the RAS procedure allocates the prediction error uniformly (and unjustifiably) over other elements. He argues, however, that in order to make use of additional information and improve the projection results, an I-O table should be decomposed into stable and unstable sub-blocks of coefficients and then *a priori* constraints be imposed only on the stable ones.

Utilizing a measure of stability, R_{ij} , suggested by Leontief (1951) as:

$$\text{(7-5)} \qquad R_{ij} = (2)\left(\frac{a_{ij}^{0} - a_{ij}^{t}}{a_{ij}^{0} + a_{ij}^{t}}\right)$$

and using this measure in the biproportional sense, i.e.,

$$\text{(7-6)} \qquad R_{ij} = (2)\left(\frac{a_{ij}^{t} - \tilde{a}_{ij}^{t}}{a_{ij}^{t} + \tilde{a}_{ij}^{t}}\right)$$

he proposes an algorithm for Modified RAS, called Extended RAS or ERAS. In his model, the constraints are imposed on these sub-blocks of stable coefficients. The procedure reveals improvement over simple RAS method.

7.2.2) Modified Lagrangian Method. In a logic similar to that of modified RAS technique, one can incorporate additional available information on elements of the target year matrix into the table estimated via the Lagrangian method. That is to say, the known (or exogenously estimated) coefficients will assume their proper (estimated) values and then the Lagrangian technique will be employed to estimate the unknown portion of the target year table. This technique has been utilized by Morrison and Thuman (1980) for 1963, 1967, and 1972 tables of Washington state. It is logically plausible and operationally feasible, hence it has been included in the current project.

As mentioned before, a shortcoming of the Lagrangian multiplier method in general is that it does not guarantee the non-negativity of the estimated coefficients. This problem, conceptually, can be removed by imposing non-negativity constraints on the solution. These constraints, however, will be in the form of inequalities, which in turn, make the problem a quadratic programming one.

In general, to solve a modified Lagrangian problem, the known elements of the target year matrix can be either incorporated into the objective function or be included in the set of constraints. In either case the results will be the same. The choice of minimand is also open to any logical one. However, the Friedlander's minimand, as indicated by the results obtained previously, performed better than other Lagrangian type minimands. Hence, in the interest of continuity and compatibility, the Friedlander minimand has been selected here for modification purposes. Formally:

(7-7) *Minimize* $$\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{(\tilde{x}_{ij}^{t}-x_{ij}^{0})^{2}}{x_{ij}^{0}}$$

S.T:

(7-8) $$\sum_{j=1}^{n}\tilde{x}_{ij}^{t}=u_{i}^{t}$$

(7-9) $$\sum_{i=1}^{n}\tilde{x}_{ij}^{t}=v_{j}^{t}$$

(7-10) $$\tilde{x}_{kl}^{t}=m_{kl}$$

where S = {k,l} is the set of ordered pairs of known elements of the target year matrix and all other terms are the same as before. The auxiliary function, then, can be written as (7-11) below:

$$\Phi=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{(\tilde{x}_{ij}^{t}-x_{ij}^{0})}{x_{ij}^{0}}-$$

$$\sum_{i=1}^{n}\lambda_{i}\left[u_{i}^{t}-\sum_{j=1}^{n}\tilde{x}_{ij}^{t}\right]+\sum_{j=1}^{n-1}\mu_{j}\left[v_{j}^{t}-\sum_{i=1}^{n}\tilde{x}_{ij}^{t}\right]+\sum_{k}\sum_{l}\theta_{kl}\left[m_{kl}-\tilde{x}_{kl}^{t}\right]$$

$$for\ \{k,l\}\in S$$

Partial derivatives of this function at optimum are (7-12), (7-13), (7-14) below for Lagrangian multipliers,

$$\frac{\partial \Phi}{\partial \lambda_i} = u_i^t - \sum_j^n \tilde{x}_{ij}^t = 0$$

$$\frac{\partial \Phi}{\partial \mu_j} = v_j^t - \sum_i^n \tilde{x}_{ij}^t = 0$$

$$\frac{\partial \Phi}{\partial \theta_{kl}} = m_{kl}^t - x_{kl}^t = 0$$

and (7-15), (7-16), and (7-17) for the I-O coefficients.

$$\frac{\partial \Phi}{\partial x_{ij}^t} = \frac{2(\tilde{x}_{ij}^t - x_{ij}^0)}{x_{ij}^0} - \lambda_i - \mu_j = 0$$

$$\frac{\partial \Phi}{\partial x_{i.n}^t} = \frac{2(\tilde{x}_{i.n}^t - x_{i.n}^0)}{x_{i.n}^0} - \lambda_i = 0$$

for $i=1,2,\ldots,n;\quad j=1,2,\ldots,n-1;\quad \{i,j\} \notin S=\{k,l\}$

$$\frac{\partial \Phi}{\partial x_{kl}^t} = \frac{2(\tilde{x}_{kl}^t - x_{kl}^0)}{x_{kl}^0} - \lambda_k - \mu_l - \theta_{kl} = 0$$

for $\{k,l\} = S$

As was shown previously, the general solution of the Lagrangian (without additional constraints) will be (7-18) and (7-19) below:

$$\tilde{x}_{ij}^{t} = x_{ij}^{0}\left[1 + \left(\frac{1}{2}\right)(\lambda_i + \mu_j)\right]$$

for the first through n-1 columns

$$\tilde{x}_{i.n}^{t} = x_{ij}^{0}\left[1 + \left(\frac{1}{2}\right)(\lambda_i)\right]$$

for the n-th column

Keeping in mind that $\sum_{j}^{n} \tilde{x}_{ij}^{t} = u_i^t$ and $\sum_{j}^{n} x_{ij}^{0} = u_i^0$, one can

rearrange and simplify these equations to get:

(7-20) $$2(u_i^t - u_i^0) = u_i^0\lambda_i + \sum_{j}^{n-1} x_{ij}^0 \mu_j$$

in column direction and:

(7-21) $$2(v_j^t - v_j^0) = v_j^0\mu_j + \sum_{i}^{n} x_{ij}^0 \lambda_i$$

in row direction.

In the general case, the simultaneous solution of this system will give the desired values of λ_i 's and μ_j 's, and thus the desired estimated values for the target year matrix. The system, however, must be adjusted to allow the utilization of known coefficients. In other words, since some of the elements of the target year matrix are known in advance, there is no need to estimate them via the updating technique, hence the system of equations must be adjusted accordingly. Adjusting the general solution for the known values in the target year involves their exclusion from the optimization process by subtracting them from the appropriate row and column marginal totals of the target year matrix, as well as subtracting the corresponding elements in the base year from their proper marginals.

Let $[M^t]$ and $[M^0]$ be two (n) by (n) matrices whose elements are known coefficients of the target year and their corresponding coefficients in the base year (i.e., m_{kl}^t 's and m_{kl}^0 's) in proper places and zeroes elsewhere. Inclusion of these two matrices in the general solution system will yield the desired results. Formally (7-22) and (7-23) below:

$$\left[(u_i^t-\sum_j m_{ij}^t)-(u_i^0-\sum_j m_{ij}^0)\right]=\frac{1}{2}(u_i^0-\sum_j m_{ij}^0)(\lambda_i)+\frac{1}{2}\sum_j^{n-1}\left[(x_{ij}^0-m_{ij}^0)(\mu_j)\right]$$

$$\left[(v_j^t-\sum_i m_{ij}^t)-(v_j^0-\sum_i m_{ij}^0)\right]=\frac{1}{2}(v_j^0-\sum_i m_{ij}^0)(\mu_j)+\frac{1}{2}\sum_i\left[(x_{ij}^0-m_{ij}^0)(\lambda_i)\right]$$

The system can be written in a more compact form as:

(7-24)
$$F_i = A_i\,\lambda_i + \sum_j^{n-1} Y_j\,\mu_j$$
$$D_j = \sum_i Y_i\,\lambda_i + B_j\,\mu_j$$

Where:

$$F_i = \left[(u_i^t - \sum_j m_{ij}^t) - (u_i^0 - \sum_j m_{ij}^0)\right]$$
$$D_j = \left[(v_j^t - \sum_i m_{ij}^t) - (v_j^0 - \sum_i m_{ij}^0)\right]$$

(7-25)
$$A_i = (\frac{1}{2})\left(u_i^0 - \sum_j m_{ij}^0\right)$$
$$Y = (\frac{1}{2})\left(x_{ij}^0 - m_{ij}^0\right)$$
$$B_j = (\frac{1}{2})\left(v_j^0 - \sum_i m_{ij}^0\right)$$

or in matrix notation:

(7-26)
$$\begin{bmatrix} F \\ D \end{bmatrix} = \begin{bmatrix} A & Y & \lambda \\ Y' & B & \mu \end{bmatrix}$$

which, as shown before, yields the values of λ and μ as:

(7-27)
$$\lambda = [A^{-1}]\ [F - Y\,\mu]$$
$$\mu = [B - Y'(A^{-1})\,Y]^{-1}\,[D - Y'(A^{-1})\,F]$$

Having the appropriate values of λ and μ , one can obtain the relevant estimates of the target year coefficients (i.e., $\tilde{x}_{ij}^{t}$'s). These estimates, in

conjunction with the known values of the target year coefficients will make up the complete technology coefficient of the I-O table for that year. To facilitate a better understanding of the system, its detailed matrix format is offered in the system (7-28) below.

$$
\begin{bmatrix} F_1 \\ F_2 \\ \cdot \\ \cdot \\ \cdot \\ F_n \\ \hline D_1 \\ D_2 \\ \cdot \\ \cdot \\ D_{n-1} \end{bmatrix}
=
\left[\begin{array}{cccccc|ccccc}
\sum_j^n S_{1j} & 0 & \cdot & \cdot & \cdot & 0 & S_{11} & S_{12} & \cdot & \cdot & S_{1,n-1} \\
0 & \sum_j^n S_{2j} & \cdot & \cdot & \cdot & 0 & S_{21} & S_{22} & \cdot & \cdot & S_{2,n-1} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \sum_j^n S_{n,j} & S_{n,1} & S_{n,2} & \cdot & \cdot & S_{n,n-1} \\
\hline
S_{11} & S_{21} & \cdot & \cdot & \cdot & S_{n,1} & \sum_i^n S_{i1} & 0 & \cdot & \cdot & 0 \\
S_{12} & S_{22} & \cdot & \cdot & \cdot & S_{n,2} & 0 & \sum_i^n S_{i2} & 0 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
S_{1,n-1} & S_{2,n-1} & \cdot & \cdot & \cdot & S_{n,n-1} & 0 & 0 & \cdot & \cdot & \sum_i^n S_{i,n-1}
\end{array}\right]
\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_n \\ \hline \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \mu_{n-1} \end{bmatrix}
$$

It should be noted that in the above system

$$u_i^t = \sum_j^n \tilde{x}_{ij}^t \;, \quad u_i^0 = \sum_j x_{ij}^0 \;, \text{ and } \left(\frac{1}{2}\right)(x_{ij}^0 - m_{ij}^0) = S_{ij} \;.$$

7.2.3) <u>**Modified NAIVE Method**</u>. This method is utilization of the most recent survey based coefficients along with some additional exogenous information (selected via a given criteria). In other words, selected coefficients of the target year matrix are estimated and the remaining coefficients are assumed to have no change during the period under study. This method is utilized to shed some lights on the merit of updating methods when some exogenous data is available. It can also serve as a test of temporal stability of I-O coefficients.

7.3) CRITERIONS FOR EXOGENOUS ESTIMATION:

Regardless of which technique one uses for updating purposes, in modifying the methods via incorporation of exogenous information, one must decide which coefficients should be estimated exogenously. That is to say, which additional information should be incorporated into the updating process in order to obtain a more accurate estimate of the target year matrix. There are two general answers to this question. First, availability of data (i.e., additional information available from outside of the I-O model), and second, selective targeting (i.e., active and systematic search within the I-O model to select cells for exogenous estimation). In other words, given some resources, researchers can direct their efforts in targeting certain cells or industries for individual estimation and inclusion in the updating process.

The identification of an appropriate subset of cells and industries, may not only help to improve estimation of the target year matrix, but also can prove beneficial in choice of sectors for aggregation purposes, i.e., some of the limited resources may be saved via aggregation of larger tables without loss of relevant accuracy, given that the appropriate subset of cells remain dissaggregated.

The identification process, however requires establishment of an appropriate scheme. Conceptually there could be many criterions for selective targeting. Three such criterions are used in the present project. This choice was due to logical coherence, practicality, widespread acceptability, and computational ease of these measures. These criterions, along with the case of available data from outside of the I-O model, are briefly explained below.

7.3.1) Availability of Data. As mentioned earlier, in most instances, the choice of which exogenous information to be used is dictated to the I-O analysts by the circumstances such as availability of data and resources. Generally, however, whatever information about the target year exist should be incorporated into the projection model. The new information can be either direct numerical value(s) for some cell(s) of the target year matrix; or some knowledge of factors such as technological changes, relative price changes, laws and regulations, product mix, and other such relevant factors. The inclusion of these information normally comes at no or minimal cost to the I-O researchers. But, as noted earlier, this inclusion, in some cases can result in deterioration of the updated matrix.

7.3.2) Key Cells or Industries. Frequently, some industries acquire a preeminent position in an economy, hence it will be of paramount importance that their estimates to be as accurate as possible. The acquisition of such a dominant position could be due to variety of reasons. It could be because of heavy reliance of the economy on that sector's output (e.g., oil for OPEC countries, coal for coal producing states, etc.); it could be as a result of some social or political situation (e.g., steel during wartime, foodstuff during droughts, etc.); it could be as a consequence of some policies (e.g., construction materials for housing projects in response to homelessness, labor intensive industries in response to unemployment, etc.); and so on. Under these circumstances, researchers may attempt to exogenously project these sector(s) and incorporate the results in the estimating procedure in the hope of obtaining a more accurate table for the target year. It should be noted that a cell or industry chosen here, may not necessarily be mathematically important to the matrix or structurally important to the table. That is to say the location of the "key" cells chosen here may not be such that changes the mathematical properties of the technology matrix (e.g., condition number, eigen value, etc.) or be structurally important to the I-O table (e.g., having the most total effects on or highest linkage to other cells). Simply put, this assertion indicates that selection of cells via "key" criterion may result in choosing cells that, by other measures, are not the most important or influential cells in the matrix. If such considerations are important to a researcher, other selection methods must be utilized.

7.3.3) Large Coefficients. In their quest to modify updating techniques and improve the estimates, many researchers realized that I-O tables are heavily dependent on a fairly small number of key coefficients, and that these coefficients often posses relative temporal stability. Results of experiments and investigations by analysts such as Paelink and Waelbroeck (1963, cited in Lecomber, 1975), Simpson and Tsukui (1965), Tilanus (1966), Bacharach (1970), Allen (1974), Almon (1974), Lecomber (1975), Allen and Lecomber (1975), Parikh (1979), Butterfield and Mules (1980), Jensen and West (1980), Hewings and Romanos (1981, cited in Kim, 1984), and Israilevich (1986) are all point out this phenomenon.

In addition, most of the results seem to indicate that the large coefficients assume this crucial role in I-O tables. Therefore, if large coefficients in I-O tables exert a heavy and significant influence on the table as a whole, an allocation of resources to determine the value of large coefficients exogenously and incorporating them in the updating scheme, may well be justified by the resultant accuracy of the updated table. That is to say, a criteria for selective targeting could be the magnitude of the coefficients in the relevant table as compared to a predetermined threshold. It should be noted that the "largeness" here is entirely arbitrary and the value of the threshold depends on the researcher's needs and resources.

7.3.4) Most Important Parameters. From the studies mentioned in the previous section, it is evident that I-O researchers agree on the heavy influence of a relatively small and fairly stable number of cells on the whole matrix. It is also clear that most researchers believe that these important cells are the largest ones in the table. Their purpose was to locate the subset of a matrix that should be subject to exogenous estimation.

In a closely related line of investigation, other researchers, attempted to decompose matrices (I-O tables) and device systematic schemes to identify the cells whose alteration could result in highest perturbation or change in the matrix.

The problem was basically twofold, "tolerance problem" and "most important parameter problem." The first one was concerned with determining the worst case boundaries (tolerance) of each element of an inverse matrix as the result of parametric uncertainties (variation of the value within a given interval) in any element of the direct coefficient matrix. The second question was to find elements of the direct matrix that, if changed (within an interval), can cause a significant (in some agreed upon sense) change in any element of the inverse matrix.

Researches by Sherman and Morrison (1950) and Berman (1953) led to a solution for the second question, and Christ (1955) presented an answer, albeit for a specific group of matrices, to the first one. Other contributors to this line of research include Chenery and Clark (1959), Simpson and Tsukui (1965), Allen (1974), and, as reported by Sebald (1974) Moore (1955), Goldstine and Von Neumann (1951), Autin and Rioux (1971), and Tomlin (1973) who took a stochastic approach to the problem.

Sebald (1974) criticized Christ's technique for containing too restrictive assumptions, the Sherman-Morrison/Berman method for being incomplete and computationally demanding, and the stochastic approach for being costly for large matrices. Then, he attempts tightening the boundaries as well as generalizing Christ's technique and proposes an algorithm and a routine to sufficiently reduce tedious and time consuming calculations required by Sherman-Morrison/Berman method. Bullard and Sebald (1977) utilize the proposed scheme to select most important 2% of the parameters in the table, and report that updating of this small subset yields substantial gain in accuracy and efficiency.

Similar results, through different approaches, are reported by Jensen and West (1980) and West (1982). According to these studies, removal of a large portion of elements of an I-O table will not significantly affect the results, and the importance of each element to the whole table, depends on magnitude and location of that particular element. Hewings (1984) and Sonis and Hewings (1989) proposed another method of identifying the most important elements, which can accommodate a variety of coefficient changes. In conformity with previous studies, they also found that in input-output tables, only small number of coefficients are what they term "inverse important" elements. Inverse importance refers to the cumulative effects (or what they refer to as the "field of influence") of a change in any given element of the direct matrix on the whole system. The issues of temporal stability of these inverse important coefficients, as well as a suitable method of measuring the gain in accuracy of the model due to utilization of exogenous information were addressed.

In another study, Bullard and Sebald (1988) try to evaluate the usage of Monte Carlo simulation to analyze the effects of stochastically perturbing an Input-Output table. They studied the relative effects of aggregation, mean value bias, magnification of input errors, etc., in order to obtain a tighter bound on error magnification and cancellation of parametric uncertainties due to opposite deviations. They found that aggregation of tables does not significantly change the output uncertainties, and attributed this to possibility of effect of smaller variances being offset by decrease in error cancellation.

The current project uses the Bullard and Sebald approach- though with some deviation - for identification of the most important parameters in its modification schemes. Therefore, a brief explanation of the approach is warranted. Some background and introduction to the "tolerance problem," however, seems appropriate to facilitate comprehension of this approach. The following explanation, in addition to the sources already mentioned in this section, draws upon Isaacson and Keller (1966), Noble (1969), Strang (1980), and Pearson (1983).

It is well known that the Euclidean length of a vector is equal to the square root of the sum of squares of its elements, i.e.,:

$$\| F \| = \sqrt{f_1^2 + f_2^2 + \ldots\ldots + f_n^2} \qquad (7\text{-}29)$$

Extending this concept to a matrix, one can think of "norm" of a matrix as a scaler which gives an overall measure of the size of that matrix. For matrix A, this can be defined as the square root of the sum of the squares of all its elements. Formally:

$$\| A \| = \sqrt{\sum_{i}^{n}\sum_{j}^{n} a_{ij}^{2}} \quad (7\text{-}30)$$

Or the square root of the largest eigenvalue (λ_{max}) of

A^TA). That is:

$$\| A \| = \sqrt{\lambda_{max}(A^TA)} \quad (7\text{-}31)$$

Now, if a matrix is perturbed multiplicatively or additively (i.e., all its elements are multiplied by a given number or a given number is added to all the elements of the matrix) and the norm of this perturbed matrix is calculated, comparison of this norm with the unperturbed norm may be used as some measure of the effects of that particular perturbation on the matrix, viz, the error amplifying power of the matrix. The magnitude of this "amplification" depends on the value of the matrix's "condition number" (defined as $C = \| A \| \, \| A^{-1} \|$). The condition number measure the sensitivity of a matrix to a given perturbation, i.e., whether the matrix is "well conditioned" (not very sensitive to a small change) or "ill-conditioned" (very sensitive to a small change). Utilizing the concepts of norm and condition number, a bound on this error amplifying power can be defined.

Consider the solution to an I-O problem:

(7-32) $$X = (I - A)^{-1} Y$$

Let δA represent a given perturbation on (I - A). Then, the tolerances on the inverse of this matrix can be specified as:

(7-33) $$\frac{\|[I - (A + \delta A)]^{-1} - [I - A]^{-1}\|}{\|[I - A]^{-1}\|} \leq \frac{C \dfrac{\|\delta A\|}{\|I - A\|}}{1 - C \dfrac{\|\delta A\|}{\|I - A\|}}$$

given that $\|[I - A]^{-1}\| . \|\delta A\| < 1$. That is to say, this equation provides boundaries on the norm of the difference between the original matrix and its perturbed inverse. The boundaries however, as noted by Sebald (1974) and Bullard and Sebald (1975) are only on the norm (not elements) and are very loose. They suggest a different procedure. The question they ask is, given known tolerances of each element in matrix [A], what are the worst case positive and negative tolerances on the elements of the corresponding Leontief inverse? Then they construct two worst case perturbation matrices for the positive and negative tolerances, i.e., matrices that can cause worst case positive and negative simultaneous increase on all elements of $(I - A)^{-1}$. The upper and lower element by element tolerances, then, are determined via inverses of these two matrices. They also demonstrate that these matrices can be constructed for many I-O tables by simply allowing all values of the [A] matrix to simultaneously assume their highest (or lowest) values, and a much tighter bounds than the norm bounds are obtained. The model, however, assumes the errors in all elements occur simultaneously and in the same direction.

Normally, an I-O table contains large number of coefficients, and tolerance intervals for each element may be obtained. But, as noted earlier, many studies suggest that only a small number of these cells can significantly influence the solution. Therefore efficiency dictates selection of the most important elements and concentration of efforts on these coefficients. Having the worst case boundaries on the elements of an inverse matrix due to a perturbation in an element of the direct matrix, a researcher now must define a criterion of importance and accordingly choose a subset of the I-O coefficients for closer scrutiny. Sebald (1974) and Bullard and Sebald (1975) define "importance" as follow.

An element, a_{ij} of a matrix A is considered as inverse important if an α percent change in its value can lead to at least a β percent (a prescribed value) in any element of the matrix $(I-A)^{-1}$. Under two special cases when elements of the [A] matrix and the Leontief inverse matrix are very small, Sebald modifies the definition of "importance." In the first, case a_{ij} is considered as important if an additive change of γ in its value cause at least a β percent change in any element of $(I-A)^{-1}$. In the second case, an element a_{ij} is said to be important if an α percent change in its value cause an element of the appropriate Leontief matrix to exceed its original value by ρ (a given amount). He then proceeds to show that only evaluation of positive changes are necessary, since the solution perturbations due to negative changes are always smaller than or equal to the perturbations resulting from positive changes of the same magnitude. Formally, Bullard and Sebald's importance function can be defined as:

(7-34) $$J = g[(I - A)^{-1}, Y]$$

Where (J) can assume a form of an scaler, a vector, or a matrix of order smaller than or equal to $(I-A)^{-1}$.

An element of [A] is considered important if a given change in its value leads to a change in any element of (J) to be greater than or equal to a predetermined threshold, i.e., a perturbation δA of any element of the original matrix and the corresponding change in the importance function ($\delta J = g[(I - A - \delta A)^{-1}, Y]$) is compared against a pre set importance threshold to determine its importance for the model. Bullard and Sebald, utilize Sherma-Morrison relationship to evaluate δJ . Specifically, the change in any element of $(I-A)^{-1}$ due to a change in any a_{ij} (i.e., δJ_{mn}) is obtained via:

$$\delta J_{mn} = \left| [(I - A - \delta A)^{-1}_{mn}] - (I - A)^{-1}_{mn} \right|$$

(7-35)

$$= [(I - A)^{-1}_{mi}] [(I - A)^{-1}_{jn}] \left| \frac{\delta a_{ij}}{1 - [(I - A)^{-1}_{ji}] [\delta a_{ij}]} \right| .$$

provided that $1 - [(I-A)^{-1}_{ji}] [\delta a_{ij}] \neq 0$. Then, a_{ij} is considered to be inverse important if α percent change in its value causes any element of $(I-A)^{-1}$ to change by at least β percent (or any element of δJ to be greater than or equal to a predetermined level). The threshold test will be:

(7-36)

$$\left[(I - A)^{-1}_{mi}\right] \left[(I - A)^{-1}_{jn}\right] \geq \left[\frac{\beta}{100}\right] \left[\frac{1 - (I - A)^{-1}_{ji}(\frac{\alpha}{100} a_{ij})}{(\frac{\alpha}{100} a_{ij})}\right] \left[(I - A)^{-1}_{mn}\right]$$

The importance of any given element with respect to the entire importance function can be obtained through:

(7-37)

$$\sum_m \sum_n \Psi_{mn}$$

where:

(7-38)

$$\Psi_{mn} = \nu = \frac{[(I - A)^{-1}_{mi}] [(I - A)^{-1}_{jn}]}{\tau} = \frac{\delta J_{mn}}{\tau}$$

for $\nu \geq 1$, τ = *applicable threshold*

$\Psi_{mn} = 0$ *otherwise*

Thus, a researcher can determine the importance threshold and through usage of this technique, determine the most important parameters, i.e., the parameters that exert the largest influence on the solution of the I-O model, and concentrate his/her efforts on the accurate estimation of these elements.

It should be noted that the identification of important parameters can be utilized, not only for selective targeting in modification of updating procedures, but also in determination of critical points in impact analysis. The identification may also be of some use in the survey phase of the construction of the table, as it can point to those coefficients that must be constructed with greater accuracy.

Application of three selection criterions (i.e., key sectors, large cells, and most important parameters) to three selected updating techniques (i.e., RAS, Lagrangian, and Naive) generates nine estimates of the target matrix, via incorporation of exogenous data.

7.4) RESIDUAL MINIMUM METHOD:

In addition to the above approaches, another modification method is used, namely the Residual Minimum Method. This method, though not incorporating exogenous information as such, still can be considered as a modified updating procedure. For it modifies, albeit somewhat mechanically, the estimated coefficients. This method, as reported by Kaneko (1983), was developed by the Economic Planning Agency in Japan. The rationale for the name is not explained in the source.

The present study utilizes an altered version of this method. The alteration was deemed necessary due to inadequacy of explanation as well as errors in mathematical representation in the source. It should be emphasized that these alterations represent the present researcher's understanding of what the method ought to be, and may not necessarily be agreed to by the original developers of the technique.

The method attempts to attach some weight to the coefficients in the updating process and accordingly, it is hoped, improve the resultant updated matrix. Formally, the residual minimum method seeks to:

$$\text{(7-39)} \qquad Minimize \sum_{i=1}^{n}\sum_{j=1}^{n}\left[\frac{(\tilde{x}_{ij}^{t} - x_{ij}^{0})^{2}}{x_{ij}^{0}}\right]\left[x_{ij}^{0}\right]^{\rho}$$

subject to:

$$\text{(7-40)} \qquad \sum_{j=1}^{n}\tilde{x}_{ij}^{t} = u_{i}^{t}$$

$$\text{(7-41)} \qquad \sum_{i=1}^{n}\tilde{x}_{ij}^{t} = v_{j}^{t}$$

where $\tilde{x}_{ij}^{t}$ represent estimated elements of the target year matrix, x^0_{ij} denote their corresponding values in the base year, ρ refers to the adjusting coefficient, and u_i^t and v_j^t are the known marginal totals of the target year matrix. In this formulation, estimated coefficients are weighted by their base year magnitude (which themselves are adjusted by some subjectively selected coefficient). The appropriate Lagrangian will be:

(7-42)

$$\Phi = \sum_{i=1}^{n}\sum_{j=1}^{n}\left[\frac{(\tilde{x}^t_{ij}-x^0_{ij})^2}{x^0_{ij}}\right][x^0_{ij}]^{\rho} + \sum_{i=1}^{n}\lambda_i\left[u_i^t - \sum_{j=1}^{n}\tilde{x}^t_{ij}\right] + \sum_{j=1}^{n-1}\mu_j\left[v_j^t - \sum_{i=1}^{n}\tilde{x}^t_{ij}\right]$$

Taking the appropriate partial derivatives and setting them equal to zero will yield:

(7-43) $$\frac{\partial\Phi}{\partial\tilde{x}^t_{ij}} = \frac{2(x^0_{ij})^{\rho}(\tilde{x}^t_{ij})}{x^0_{ij}} - \frac{2(x^0_{ij})(x^0_{ij})^{\rho}}{x^0_{ij}} - \lambda_i - \mu_j = 0$$

for i = 1, 2, ..., n and j = 1, 2, ..., n-1

and:

(7-44) $$\frac{\partial\Phi}{\partial\tilde{x}^t_{i.n}} = \frac{2(x^0_{i.n})^{\rho}(\tilde{x}^t_{i.n})}{x^0_{i.n}} - \frac{2(x^0_{i.n})(x^0_{i.n})^{\rho}}{x^0_{i.n}} - \lambda_i = 0$$

for i = 1, 2, ..., n and j = n

(7-45) $$\frac{\partial\Phi}{\partial\lambda_i} = u_i^t - \sum_{j=1}^{n} \tilde{x}_{ij}^t = 0$$

and the followings for the Lagrangian multipliers:

(7-46) $$\frac{\partial\Phi}{\partial\mu_j} = v_j^t - \sum_{i=1}^{n} \tilde{x}_{ij}^t = 0$$

From the first two equations, one can get:

(7-47) $$\tilde{x}_{ij}^t = \frac{2\,(x_{ij}^0)^{\rho+1} + \lambda_i\, x_{ij}^0 + \mu_j\, x_{ij}^0}{2\,(x_{ij}^0)^{\rho}}$$

for the first to (n-1)-th columns, and

(7-48) $$\tilde{x}_{i.n}^t = \frac{2\,(x_{i.n}^0)^{\rho+1} + \lambda_i\, x_{i.n}^0}{2\,(x_{i.n}^0)^{\rho}}$$

for the n-th column.

Substituting in the other two equations, rearranging, and simplifying gives:

(7-49) $$u_i^t - \sum_{j=1}^{n} x_{ij}^0 = \frac{1}{2}\sum_{j=1}^{n} \left[(\lambda_i)\,(x_{ij}^0)^{1-\rho}\right] + \frac{1}{2}\sum_{j=1}^{n-1} \left[(\mu_j)\,(x_{ij}^0)^{1-\rho}\right]$$

and:

(7-50) $$v_j^t - \sum_{i=1}^{n} x_{ij}^0 = \frac{1}{2}\sum_{i=1}^{n} \left[(\lambda_i)\,(x_{ij}^0)^{1-\rho}\right] + \frac{1}{2}\sum_{i=1}^{n} (\mu_j)\,(x_{ij}^0)^{1-\rho} \; .$$

Let:

$$(7\text{-}51)\qquad \begin{aligned} F_i &= u_i^t - \sum_{j=1}^{n} x_{ij}^0 \\ D_j &= V_j^t - \sum_{i=1}^{n} x_{ij}^0 \\ A_i &= \left(\frac{1}{2}\right) \sum_{j=1}^{n} (x_{ij}^0)^{1-\rho} \\ B_j &= \left(\frac{1}{2}\right) \sum_{i=1}^{n} (x_{ij}^0)^{1-\rho} \\ Y &= \left(\frac{1}{2}\right) (x_{ij}^0)^{1-\rho} \end{aligned}$$

Then, one can get:

$$(7\text{-}52)\qquad \begin{aligned} F_i &= A_i\lambda_i + \sum_{j=1}^{n-1} Y_j\mu_j \\ D_j &= \sum_{i=1}^{n} Y_i\lambda_i + \mu_j B_j \end{aligned}$$

Simultaneous solution of this system will give the desired values for the Lagrangian multipliers μ and λ ,

which can be used to obtain estimates of the target year matrix through:

$$(7\text{-}53)\qquad x_{ij}^t = (x_{ij}^0)\left[1 + \left(\frac{1}{2}\right)(x_{ij}^0)^{-\rho}(\lambda_i + \mu_j)\right]$$

for: 1*st through* (n-1)-*th sector*

$$(7\text{-}54)\qquad x_{i.n}^t = (x_{i.n}^0)\left[1 + \left(\frac{1}{2}\right)(x_{i.n}^0)^{-\rho}(\lambda_i)\right]$$

for the n-th sector.

It should be noted that, here, the estimated elements of the target year table are weighted by the adjusted (by ρ) value of their corresponding elements in the base year.

The solution in matrix notation can be presented as:

$$\text{(7-55)} \qquad \begin{bmatrix} F \\ D \end{bmatrix} = \begin{bmatrix} A & Y \\ Y' & B \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix}$$

which will yield:

$$\text{(7-56)} \qquad \begin{aligned} \lambda &= (A)^{-1}[F - Y\mu] \\ \mu &= [B - Y'(A)^{-1}Y]^{-1}[D - Y'(A)^{-1}F] \end{aligned}$$

It is interesting to note that if the value of adjusting coefficient ρ is set equal to one, the residual minimum method will be identical to the Friedlander approach. Value of zero for this adjusting coefficient makes the method identical to that of Almon.

A more detailed and expanded matrix presentation of the system is presented below for explicative purposes:

$$
\begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \\ \hline D_1 \\ D_2 \\ \vdots \\ D_{n-1} \end{bmatrix}
=
\left[\begin{array}{cccc|cccc}
\frac{1}{2}\sum_{j}^{n}(x_{1j}^{0})^{1-\rho} & \cdots & \cdots & 0 & \frac{1}{2}(x_{11}^{0})^{1-\rho} & \frac{1}{2}(x_{12}^{0})^{1-\rho} & \cdots & \frac{1}{2}(x_{1,n-1}^{0})^{1-\rho} \\
0 & \frac{1}{2}\sum_{j}^{n}(x_{2j}^{0})^{1-\rho} & \cdots & 0 & \frac{1}{2}(x_{21}^{0})^{1-\rho} & \frac{1}{2}(x_{22}^{0})^{1-\rho} & \cdots & \frac{1}{2}(x_{2,n-1}^{0})^{1-\rho} \\
\vdots & & & \vdots & \vdots & & & \vdots \\
0 & \cdots & \cdots & \frac{1}{2}\sum_{j}^{n}(x_{nj}^{0})^{1-\rho} & \frac{1}{2}(x_{n,1}^{0})^{1-\rho} & \cdots & \cdots & \frac{1}{2}(x_{n,n-1}^{0})^{1-\rho} \\
\hline
\frac{1}{2}(x_{11}^{0})^{1-\rho} & \cdots & \cdots & \frac{1}{2}(x_{n,1}^{0})^{1-\rho} & \frac{1}{2}\sum_{i}^{n}(x_{i1}^{0})^{1-\rho} & \cdots & \cdots & 0 \\
\frac{1}{2}(x_{12}^{0})^{1-\rho} & \frac{1}{2}(x_{22}^{0})^{1-\rho} & \cdots & \cdots & 0 & \frac{1}{2}\sum_{i}^{n}(x_{i,2}^{0})^{1-\rho} & \cdots & 0 \\
\vdots & & & \vdots & \vdots & & & \vdots \\
\frac{1}{2}(x_{1,n-1}^{0})^{1-\rho} & \cdots & \frac{1}{2}(x_{n.n-1}^{0})^{1-\rho} & & 0 & \cdots & \cdots & \frac{1}{2}\sum_{i}^{n}(x_{i,n-1}^{0})^{1-\rho}
\end{array}\right]
\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \hline \mu_1 \\ \mu_2 \\ \vdots \\ \mu_{n-1} \end{bmatrix}
$$

The afformentioned methods are utilized as the basis of ten modified experiments, results of which are reported in the next chapter.

CHAPTER EIGHT

RESULTS OF MODIFICATION THROUGH INCORPORATION OF EXOGENOUS INFORMATION

"Give me a fruitful error any time, full of seeds, bursting with its own corrections. You can keep your sterile truth for yourself."

Vilferdo Pareto

8.1) INTRODUCTION:

Several issues in updating of I-O tables were raised in previous chapters. Treatment of these issues, however, were deferred to a later part of the study. The current chapter is an attempt to address some of the most important of those issues.

It was mentioned that in many instances, I-O researchers find it convenient or necessary to incorporate additional exogenous information into their estimation efforts. This can be due to reasons such as availability of data and resources, special importance of a subset of coefficients, special purpose I-O table, desired improvement in accuracy, etc. It was also argued that certain portions of I-O coefficients demonstrate more stability over time than others, hence require special treatment in the updating process. The estimates, are believed, can improve if appropriate additional information and proper methods were utilized. The search for such subsets of coefficients as well as methods were conducted in chapter seven, and several techniques of identification and incorporation of exogenous information were explored.

Based on the search, it was concluded that three categories of coefficients might demand special treatment more than other coefficients, hence should be subject of "selective targeting." These categories are "key coefficients," "large coefficients," and "most important parameters." Detailed discussion of these categories were given in chapter seven and will not be repeated here.

Next step was to select updating methods that should be utilized in conjunction with the selected targets. Based on the results reported in chapter six, and in the interest of continuity and compatibility, The top two performer among the estimation methods were selected for this purpose. These methods are RAS and FRIED. Of course it must be mentioned that technically speaking RASLAG method actually tied RAS method in ranking, hence was superior to the FRIED method. However, due to reasons that was explained earlier, the results obtained via RASLAG is identical to those of RAS. Therefore only one of these methods should qualify for further analysis. RASLAG method includes RAS as a part of the procedure, which makes the latter more appropriate candidate for additional investigation. Thus, through elimination of RASLAG method, the next best technique - that is FRIED - was chosen and included in the process.

A third method, namely NAIVE, was added to this mix, even though this method was not a top performer. There are several reasons for this inclusion. Chief among these reasons are first, the fact that it is always interesting to investigate temporal stability of I-O coefficients; and second, to the extent that the coefficients do change, what is the distribution of these changes. In other words, researchers are always interested to know the extent of the changes in I-O coefficients and whether these changes occur uniformly across the table or only on a subset of the coefficients. For if the latter is true, then limited resources can be channeled into exogenous estimation of the unstable subset and the obtained results be used in conjunction with the rest of the table. This, in turn, may make the NAIVE method a viable alternative. Inclusion of the

NAIVE method, not only allows explorations along these lines, but also provides a basis of comparison for evaluation of performances of the chosen estimation procedures. Thus, inclusion of the NAIVE method in the experiment.

Three sets of selected exogenous data were incorporated into the three updating methods, leading to nine estimation of the benchmark table. The three sets of selected exogenous data are "key," "big," and "most important" coefficients. The logic, methods, and process of these selections were discussed earlier in chapter seven and do not merit duplication here. Each set contains 497 elements, which were selected via utilization of those procedures, and constitutes one variant of each updating method. These variants or modified estimated tables are designated in this study as MDNAVKEY, MDNAVBIG, MDNAVMIP, MDLAGKEY, MDLAGBIG, MDLAGMIP, MDRASKEY, MDRASBIG, and MDRASMIP.

In addition to these nine cases, a separate method known as Residual Minimum Method (see chapter seven) was utilized and included in the comparison. This method does not use exogenous data and basically relies on a mechanical adjusting process. The estimated table is designated in this project as RESMIN and it is included due to its compatibility to other techniques utilized here, as well as its potential for yielding an improved updated table.

8.2) EVALUATION OF THE RESULTS:

All ten estimates are tested according to the same sets of tests and criterions and the results are provided below. The comparisons are carried out, as before, for both direct and inverse coefficients. The more holistic comparisons of multipliers and total outputs were left out for the same reasons as before. Excluded also, is the elaborations on points that have been already addressed in chapter six.

8.2.1) Negative Coefficients. Some of the methods may generate negative coefficients in the updating process. Considering the fact that negative numbers are unacceptable for I-O coefficients, their presence reduces the power of the respective estimation methods. If the number of such negative coefficients is large enough, the whole method will be dismissed on this ground alone.

Table 8-1 provides the number of negative coefficients generated by each of the modified updating techniques for both direct and inverse coefficients. It is evident that, with three exceptions, none of the techniques utilized here produce any negative coefficient in either direct or inverse cases. The exceptions are MDLAGKEY with three negatives for direct coefficients, MDLAGMIP with four negatives for direct coefficients and MDLAGBIG with twenty-two and four for direct and inverse coefficients respectively. Comparing

TABLE 8-1

NUMBER OF NEGATIVE COEFFICIENTS GENERATED BY EACH TECHNIQUE

TECHNIQUE	DIRECT	INVERSE
MDNAVKEY	0	0
MDNAVBIG	0	0
MDNAVMIP	0	0
MDLAGKEY	3	0
MDLAGBIG	22	4
MDLAGMIP	10	0
MDRASKEY	0	0
MDRASBIG	0	0
MDRASMIP	0	0
RESMIN	0	0

these results with the unmodified results reveals that the number of generated negative coefficients has been increased. This, however, does not seem to have any connection to the modification process and appears to be merely coincidental.

Generating negative numbers will certainly damage the viability of these methods. The extent of the damage, obviously, depends on the magnitude of the negative cells in the actual table as well as their importance within the entire table. However, mere existence (or possibility of existing) of negative coefficients may seriously jeopardize the feasibility of these updating methods.

8.2.2) <u>Non-parametric Test</u>. The Wilcoxon rank-sum test was performed on column by column basis for both direct and inverse updated matrices to verify the existence of consistent overestimation or underestimation of these columns as compared to the actual data. The null hypothesis of equality of the medians of each estimated column and their counterparts in the actual matrix were tested at 5% level against the alternative hypothesis of significant difference among the two columns. Figures A-19 through A-38 of Appendix (A) visually depict the results.

In the case of direct coefficients, the null hypothesis can be rejected for MDLAGKEY, MDRASKEY, and RESMIN for two columns each. These columns, for all three cases, represent one overestimation and one underestimation. For MDLAGMIP the null hypothesis is rejected for three columns, all of which are underestimation. In all other cases the null hypothesis is rejected once per each method, all representing an underestimation. It should be noted that same sector(s) are underestimated or overestimated by all methods, which could be an indication of some structural changes in those particular sectors.

In the case of inverse coefficients the null hypothesis can be rejected only in one case, MDNAVBIG, representing an underestimation of a column. In all other cases, one will fail to reject the null hypothesis.

Thus, according to Wilcoxon test, it may be deduced that there exist no consistent overestimation or underestimation of columns of updated matrices in either direct or inverse cases. The results are also compatible with those of unmodified cases, indicating no material change due to modification.

The ranking of the methods for direct coefficients is MDRASBIG & MDLAGBIG, MDRASKEY & MDRASMIP, MDNAVKEY, MDLAGKEY & MDLAGMIP, MDNAVBIG, MDNAVMIP, and RESMIN. The ranking for the inverse case is MDRASMIP & MDRASBIG & MDLAGBIG, MDLAGKEY & MDLAGMIP & MDRASKEY & RESMIN, MDNAVBIG & MDNAVMIP, and MDNAVKEY.

8.2.3) **Coefficient of Equality**. Examination of table 8-2 reveals that at 5% level none of the techniques predict a great number of direct coefficients accurately. The range for number of coefficients estimated within 5% of the actual values, as percentage of total coefficients, is from low of 15.7% for RESMIN to high of 24.8% for MDRASMIP, with ranking of MDRASMIP, MDRASBIG & MDLAGMIP, MDNAVMIP, MDLAGBIG, MDRASKEY, MDNAVBIG, MDNAVKEY, MDLAGKEY, and RESMIN.

The same low-high percentages, as acceptance interval widens to 10% and 20%, are 21.64% to 29.52% and 31.78% to

TABLE 8-2

SUMMARY PERFORMANCE: MODIFIED DIRECT ESTIMATES

METHOD	(1 - θ)			(θ)		
	5%	10%	20%	MEAN	SD	MAX
ACTUAL	5041	5041	5041	0.8972	0.304	1.00
MDNAVKEY	1208	1450	1942	1.6772	13.862	743.96
MDNAVBIG	1215	1413	1844	2.1333	21.729	1073.40
MDNAVMIP	1228	1448	1901	1.9540	19.675	1073.40
MDLAGKEY	1200	1454	1950	1.5753	10.057	410.85
MDLAGBIG	1225	1434	1803	2.1258	17.142	796.26
MDLAGMIP	1244	1467	1902	1.9679	15.801	857.66
MDRASKEY	1220	1469	1985	1.5783	9.880	376.99
MDRASBIG	1244	1449	1856	2.2495	19.268	835.15
MDRASMIP	1250	1488	1922	2.0291	16.256	863.41
RESMIN	792	1091	1602	1.8321	14.598	616.70

* θ = coefficient of equality and 1 - θ = degree of approximation.

39.38% respectively. The pertinent rankings are MDRASMIP, MDRASKEY, MDLAGMIP, MDLAGKEY, MDNAVKEY, MDRASBIG, MDNAVMIP, MDLAGBIG, MDNAVBIG, and RESMIN for 10% level and MDRASKEY,

MDLAGKEY, MDNAVKEY, MDRASMIP, MDLAGMIP, MDNAVMIP, MDRASBIG, MDNAVBIG, MDLAGBIG, RESMIN for 20% level.

It worth noting that in relative terms all these methods performed less favorably here than they did in the case of no exogenous information. To see this, one should remember that in modification process, 497 actual coefficients from the benchmark table were selected and inserted in the updated table (the rest of the coefficients in the updated table were estimated via various updating techniques). This implies that in each estimated table there are at least 497 cells which are identical to the value of their counterparts in the benchmark table. Then, in evaluation of each method by any given criterion, say number of cells estimated within 5% of their true value, one should discount the results to allow for this "artificially injected accuracy."

Having this in mind and comparing table 8-2 with table 6-2, one can observe that all methods actually predicted less number of coefficients within the desired level in their modified versions. This, of course, does not mean that the modified tables are less accurate than their simple unmodified counterparts. Rather, it only points out that when some part of a table is exogenously determined, there may be loss of accuracy in estimation of remaining coefficients. Hence, in deciding to incorporate exogenously determined data in a table, one should weigh the benefits of such modification against the cost of collecting additional information *and* the loss of accuracy in the estimation of remaining coefficients.

It is also noteworthy that as the acceptance interval changes, there is no stability or pattern in the ranking of the estimation methods. In modified cases, just like the original situation, the NAIVE method performs very close to other techniques. This may be an indication of stability of some of the coefficients over time. Now, if one remembers that updating procedure mechanically and uniformly adjust all base year coefficients to arrive at the updated table, an interesting question will arise. Are these stable coefficients included in the number of coefficients accurately estimated by a given method or not? That is to say, MDRASMIP-for instance- predicted 1250 cells within 5% of their true value and MDNAVMIP predicted 1228 cells within the same interval; are these 1228 cells included in the 1250 accurately estimated by MDRASMIP or not? For if the answer is yes, then one may question the justification of the whole effort to obtain accuracy an additional 22 cells. If the answer is no, the question will be the wisdom of exchanging accuracy of some cells for others. Pursuit of this line, of course, is outside of the stated purposes of the current project and will not be followed here. The point was just raised as an interesting observation and possible line of research elsewhere.

With regards to the mean and standard deviation of the estimates, it can be seen that none of the methods have a mean or standard deviation anywhere close to that of the actual matrix. The large standard errors are signs of vast variation in the estimates and large means may be due to large overestimation of the coefficients. This last point is reaffirmed by the large magnitude of maximum values of the coefficients of equality, which shows the magnification of the estimates as compared with the actual coefficients. Large standard deviations and maximum values of coefficients of equality may also be due to existence of some very small coefficients in the original data.

The ranking of the methods are MDLAGKEY, MDRASKEY, MDNAVKEY, RESMIN, MDNAVMIP, MDLAGMIP, MDRASMIP, MDLAGBIG, MDNAVBIG, and MDRASBIG based on the mean of coefficients of equality and MDRASKEY, MDLAGKEY, MDNAVKEY, RESMIN, MDLAGMIP, MDRASMIP, MDLAGBIG, MDRASBIG, MDNAVMIP, and MDNAVBIG according to the standard deviation of coefficients of equality. Consideration of maximum value of the coefficients of equality will yield a ranking of MDRASKEY, MDLAGKEY, RESMIN, MDNAVKEY, MDLAGBIG, MDRASBIG, MDLAGMIP, MDRASMIP, and MDNAVMIP & MDNAVBIG. As it can be seen, these values are high, but some of the estimation methods can limit the variation much better than the NAIVE method.

Table 8-3 summarizes the results for inverse coefficients. Here, just as the case was for unmodified

TABLE 8-3

SUMMARY PERFORMANCE: MODIFIED INVERSE ESTIMATES

METHOD	(1 - θ)			(θ)		
	5%	10%	20%	MEAN	SD	MAX
ACTUAL	5041	5041	5041	1.0000	0.0000	1.00
MDNAVKEY	734	1283	2162	1.0329	1.0071	48.28
MDNAVBIG	932	1550	2486	0.9384	0.8813	47.50
MDNAVMIP	760	1472	2550	0.8775	0.3703	7.33
MDLAGKEY	1049	1668	2670	1.0589	0.8411	38.69
MDLAGBIG	1347	2061	3035	1.0835	0.7125	25.54
MDLAGMIP	1380	2159	3243	1.0083	0.4360	15.86
MDRASKEY	1057	1689	2708	1.0561	0.8136	35.66
MDRASBIG	1367	2101	3068	1.0865	0.7275	26.33
MDRASMIP	1384	2176	3280	1.0056	0.4248	15.31
RESMIN	725	1343	2468	1.0284	0.7583	30.99

* θ = coefficient of equality and (1 - θ) = degree of approximation.

results, the estimates are mixed. That is, they are sometimes less accurate than the estimates of direct coefficients and sometimes more accurate. However, as the acceptance intervals increases, the inverse estimates tend to be more accurate. The means and standard deviations are closer to their expected values, and the magnitude of the maximum values of coefficients of equality are smaller than in the cases of direct coefficients. This fact indicates that there are less variations and magnifications in the inverse cases than there are in the direct cases.

Table 8-3 reveals that none of the methods predict very large number of coefficients within the accepted intervals. At 5% level, the percentage of coefficients accurately estimated, range from 14.38% to 27.45%. These ranges for 10% and 20% acceptance intervals are 25.45%-43.17% and 42.89%-65.07% respectively. Which means, if one can tolerate 20% variation in the value of coefficient, then about two-third of the coefficients can be estimated within this range.

The ranking, as acceptance interval increases, is rather stable. For 5% and 10% the result is identical as MDRASMIP, MDLAGMIP, MDRASBIG, MDLAGBIG, MDRASKEY, MDLAGKEY, MDNAVBIG, MDNAVMIP, MDNAVKEY, and RESMIN. In the case of 20%, the ranking is the same except that MDNAVBIG and MDNAVMIP switch their places in ranking, as do RESMIN and MDNAVKEY.

Here, again, the matter of 497 exogenously determined cells must be remembered. Specifically, "BIG" and "MIP" variants of RAS and FRIED, at all three levels, and "BIG" version of NAIVE method at 20% interval, demonstrate improvement over unmodified versions even after taking into account the 497 exogenously determined coefficients. However, adjusting for the 497 cells, the other cases perform less favorably than their unmodified counterparts. The consequences of this fact are the same as they were for direct coefficients discussed above.

The ranking based on mean of coefficients of equality is MDRASMIP, MDLAGMIP, RESMIN, MDNAVKEY, MDRASKEY, MDLAGKEY, MDNAVBIG, MDLAGBIG, MDRASBIG, and MDNAVMIP. The standard deviations and maximum values of coefficients of equality yield identical ranking as MDNAVBIG, MDRASMIP, MDLAGMIP, MDLAGBIG, MDRASBIG, RESMIN, MDRASKEY, MDLAGKEY, MDNAVBIG, and MDNAVKEY.

The frequency distributions of coefficients of equality for direct and inverse coefficients are presented in tables B-19 through B-38 of the Appendix (B), and graphically depicted in figures E-20 through E-40 of Appendix (E).

8.2.4) Regression Analysis. Detail results of regression of estimated and actual coefficients for both direct and inverse cases are provided in tables D-19 through table D-38 of Appendix (D). The summary results of major statistics related to these cases are presented in tables 8-4 and 8-5 below.

In the case of direct coefficients, all methods have

TABLE 8-4

SUMMARY RESULTS OF REGRESSION ANALYSIS, DIRECT COEFFICIENTS

Method	Intercept	t-value	Slope	t-value	R^2
MDNAVKEY	-0.000239	-2.05	1.008557	2.60	0.949
	(0.000117)		(0.003290)		
MDNAVBIG	-0.000316	-8.34	0.995724	-3.96	0.994
	(0.000038)		(0.001070)		
MDNAVMIP	-0.000171	-2.70	0.975719	-13.55	0.983
	(0.000064)		(0.001792)		
MDLAGKEY	-0.000075	-0.76	1.009915	3.59	0.964
	(0.000098)		(0.002763)		
MDLAGBIG	0.000030	0.80	0.996065	-3.77	0.995
	(0.000037)		(0.001043)		
MDLAGMIP	0.000024	0.38	0.996766	-1.78	0.984
	(0.000064)		(0.001818)		
MDRASKEY	-0.000052	-0.55	1.006880	2.58	0.966
	(0.000094)		(0.002663)		
MDRASBIG	0.000030	0.81	0.996051	-3.83	0.995
	(0.000037)		(0.001031)		
MDRASMIP	0.000021	0.34	0.997187	-1.60	0.985
	(0.000062)		(0.001761)		
RESMIN	-0.000101	-0.95	1.013344	4.48	0.958
	(0.000106)		(0.002978)		

* Figures in parenthesis are standard deviations.

TABLE 8-5

SUMMARY RESULTS OF REGRESSION ANALYSIS, INVERSE COEFFICIENTS

Method	Intercept	t-value	Slope	t-value	R^2
MDNAVKEY	-0.000570	-2.95	0.986860	-9.96	0.911
	(0.000194)		(0.001319)		
MDNAVBIG	-0.001114	-18.47	0.997142	-6.95	0.999
	(0.000060)		(0.000411)		
MDNAVMIP	-0.001072	-10.21	0.982797	-24.03	0.997
	(0.000105)		(0.000716)		
MDLAGKEY	-0.000031	-0.21	1.001094	1.11	0.995
	(0.000145)		(0.000988)		
MDLAGBIG	0.000042	0.83	0.998888	-3.17	0.999
	(0.000051)		(0.000351)		
MDLAGMIP	-0.000077	-0.83	1.001039	1.63	0.998
	(0.000093)		(0.000636)		
MDRASKEY	-0.000061	-0.44	1.002184	2.29	0.995
	(0.000140)		(0.000952)		
MDRASBIG	0.000046	0.89	0.998908	-3.13	0.999
	(0.000051)		(0.000349)		
MDRASMIP	-0.000098	-1.10	1.001956	3.21	0.998
	(0.000089)		(0.000609)		
RESMIN	-0.000170	-1.05	1.005340	4.84	0.994
	(0.000162)		(0.001103)		

* Figures in parentheses are standard deviations.

very high R^2's, exhibiting a high degree of fitness between the estimated and actual coefficients. It is possible, however, that this high coefficients of determination are due to dependence of error terms to one another. To test this possibility, the Durbin-Watson test is conducted at 1% and 5% levels. In all cases one will fail to reject the null hypothesis of no serial correlation. Hence, an absence of autocorrelation, and strong relationship between the estimated and actual coefficients may be inferred.

The ranking of the methods based on coefficients of determination is MDRASBIG & MDLAGBIG, MDNAVBIG, MDRASMIP, MDLAGMIP, MDNAVMIP, MDRASKEY, MDLAGKEY, RESMIN, and MDNAVKEY.

The regression equations in the case of inverse coefficients also display high values of R^2. Based on Durbin-Watson test at 1% and 5% levels, with some exception, the high values of coefficients of determination are not due to serial correlation. The exceptions are MDNAVKEY and MDRASMIP at 5% level and MDNAVBIG and MDNAVMIP at both 1% and 5% levels, where one will reject the null hypothesis of no serial correlation. The ranking of the methods in the inverse case, based on R^2 will be MDRASBIG, & MDLAGBIG & MDNAVBIG, MDLAGMIP & MDRASMIP, MDNAVMIP, MDRASKEY & MDLAGKEY, RESMIN, and MDNAVKEY.

The standard F-test at 1% and 0.5% levels was conducted to test the overall significance of the regression equations. As can be verified from tables F-19 through F-38 of Appendix (F), all regression lines for direct and inverse coefficients are statistically significant, since their respective F-values far surpasses the critical values of 6.63 and 7.88. Thus, for all methods, in both direct and inverse cases, one can reject the null hypothesis of no linear relationship among estimated and actual coefficients in favor the alternative hypothesis of linear relationship among them. The rankings of the techniques are identical in both direct and inverse cases as MDRASBIG, MDLAGBIG, MDNAVBIG, MDRASMIP, MDLAGMIP, MDNAVMIP, MDRASKEY, MDLAGKEY, RESMIN, and MDNAVKEY.

Testing the null hypothesis of $\alpha = 0$, at 5% level, reveals that for both direct and inverse coefficients one will fail to reject the null hypothesis in all cases except MDNAVKEY, MDNAVBIG, and MDNAVMIP. At 1% level the situation is the same except that in case of direct coefficients the null hypothesis no longer can be rejected for MDNAVKEY.

These results indicate that all methods, excluding the ones just mentioned, for both direct and inverse cases produce regression lines with an intercept value which, in accordance with a *priori* expectation, is statistically not different from zero. That is to say, good fits may have been obtained via these estimates.

The ranking based on the intercept's t-value is MDRASMIP, MDLAGMIP, MDRASKEY, MDLAGKEY, MDRASBIG, MDLAGBIG, RESMIN, MDNAVKEY, MDNAVMIP, and MDNAVBIG for the direct coefficients and MDLAGKEY, MDRASKEY, MDLAGMIP, MDLAGBIG, MDRASBIG, RESMIN, MDRASMIP, MDNAVKEY, MDNAVMIP, and MDNAVBIG for inverse coefficients.

The results here, are generally compatible with those of unmodified experiments. It worth mentioning that in all cases of inverse and direct estimates (except inverse MDNAVKEY) the value of R^2 improves with modification. It should also be noted that in modified cases generally (except for inverse coefficients MDNAVKEY and direct coefficients MDRASMIP) the t-values were larger than the unmodified cases, which implies worsening of the regression equations' fit. This however caused reversal of results only in the case of direct coefficients for three variants of NAIVE method.

In the case of slope coefficients, as mentioned before, one wants to test whether the value of $\beta = 1$ or not. Hence the t-value for this purpose can not be obtained directly from the regression results and must be calculated in the manner explained in chapter six. These values for modified estimation techniques are calculated and reported in tables 8-4 and 8-5 above. The null hypothesis of $\beta = 1$ was tested at 5% and 1% for both direct and inverse coefficients.

Evaluation of tables 8-4 and 8-5 reveals that at 5% (critical t-value of 1.96) and for direct coefficients, one can reject the null hypothesis for all methods except

MDLAGMIP and MDRASMIP. At 1% level (critical value of 2.58), MDRASKEY will be added to this list. This implies, except in these instances, none of the methods generated a slope coefficient that matches the *a priori* expected value of one. For inverse coefficients, the null hypothesis, at 5%, can be rejected for all methods except MDLAGKEY, and MDLAGMIP. At 1%, only MDRASKEY can be added to the list and in all other cases one can still reject the null hypothesis of $\beta = 1$. The implications again being that estimates, except in the three cases mentioned, do not correspond to their expected value of one.

The ranking on the basis of t-values will be MDRASMIP, MDLAGMIP, MDRASKEY, MDNAVKEY, MDLAGKEY, MDLAGBIG, MDRASBIG, MDNAVBIG, RESMIN, and MDNAVMIP for direct coefficients and MDLAGKEY, MDLAGMIP, MDRASKEY, MDRASBIG, MDLAGBIG, MDRASMIP, RESMIN, MDNAVBIG, MDNAVKEY, and MDNAVMIP for inverse coefficients.

Comparing these results with those obtained through unmodified cases shows a mix movement of t-values and results. In the first set of experiments, one could not reject the null hypothesis for RAS and FRIED methods. In the modified versions and at 1% level, however, this is true only for one variant of these methods (MDRASMIP, and MDLAGMIP) and the null hypothesis for other two variants of RAS and FRIED methods can be rejected. If the significance level is changed to 5%, one more variant of RAS, namely

MDRASKEY, will survive, and other results remain unchanged.

Same comparison for the inverse coefficients reveals that under unmodified versions the null hypothesis could not be rejected only for RAS method. In modified version, at 1% level, one variant of RAS (i.e., MDRASKEY) and two variants of FRIED (i.e., MDLAGMIP, and MDLAGKEY) will have the desired slope coefficients, and if the level of significance is changed to 5%, the RAS variant will not survive the test.

Therefore, it can be said that modification in all direct coefficient cases (except MDRASMIP, MDNAVKEY, and MDNAVBIG) causes the t-values to increase, which implies deterioration of the regression equations' fit. The same is also true for inverse cases. The only difference is the exception cases, which for inverse coefficients are MDNAVKEY, MDNAVBIG, MDLAGKEY, and MDLAGMIP.

The joint F-test, for testing the null hypothesis of simultaneous satisfaction of $\alpha = 0$ and $\beta = 1$, against the alternative hypothesis that at least one of these conditions is not satisfied, was conducted at 1% and 0.5% levels of significance. These F-values were calculated from the regression data via employment of equation 5-6, and the calculated values are presented in table 8-6 below.

It is evident from the table 8-6, that at 1% level (critical F value of 4.61) for direct coefficients, one only fails to reject the null hypothesis in the case of MDLAGMIP. If the level of significance is chosen to be 0.5% (critical

TABLE 8-6

CALCULATED F-STATISTIC

DIRECT AND INVERSE COEFFICIENTS

ESTIMATION TECHNIQUE	F-STATISTIC DIRECT	INVERSE
MDNAVKEY	5.57	63.17
MDNAVBIG	46.80	244.52
MDNAVMIP	105.15	409.14
MDLAGKEY	5.19	1.33
MDLAGBIG	19.09	17.72
MDLAGMIP	4.24	1.55
MDRASKEY	6.77	5.00
MDRASBIG	42.52	1.58
MDRASMIP	10.34	7.92
RESMIN	5.49	13.93

F-value of 5.30), MDLAGKEY will be added to this list. In other words, only these two estimates meet the *a priori* expectations. Other estimates, do not seem to have regression equations with a good fit, although at 0.5% level MDNAVKEY and RESMIN were barely failed to meet the criteria. This observation implies absence of statistical similarity between the actual and estimated matrices in the relevant cases.

For inverse coefficients at 1% level, the null hypothesis can be rejected for all cases except MDLAGKEY, MDLAGMIP, and MDRASBIG. At 0.5% level, MDRASKEY will be added to the list. Thus, the results again indicate, except for the cases mentioned, absence of statistical similarity between the estimated and actual inverse matrices.

An interesting observation can be made by comparing tables 8-6 and 6-6. That is, under modified scheme and as measured by the joint F-test criterion, the FRIED method improves rather substantially while RAS method deteriorates a good deal. The ranking of the techniques based on joint F-test for direct coefficients is MDLAGMIP, MDLAGKEY, RESMIN, MDNAVKEY, MDRASKEY, MDRASMIP, MDLAGBIG, MDRASBIG, MDNAVBIG, and MDNAVMIP. For the inverse coefficients, the ranking will be MDLAGKEY, MDLAGMIP, MDRASBIG, MDRASKEY, MDRASMIP, RESMIN, MDLAGBIG, MDNAVKEY, MDNAVBIG, and MDNAVMIP.

All other explanatory points raised in chapter six are applicable here as well and will not be restated.

8.2.5) <u>Chi-Square Test</u>. The χ^2 test was conducted between modified estimated matrices and the actual matrix to investigate the existence of similarity of distributions among estimated and actual coefficients. The test was, as explained earlier, performed at 1% and 0.5% for two sets of data, one class and forty classes. The relevant critical values are 62.248 and 65.476 for forty classes scheme and

TABLE 8-7

CHI SQUARE TEST, 40 CLASSES

	DIRECT COEFFICIENTS	INVERSE COEFFICIENT
MDNAVKEY	255.76834	34.672599
MDNAVBIG	276.30860	9.231417
MDNAVMIP	284.56575	19.519233
MDLAGKEY	64.645990	23.095880
MDLAGBIG	0.815954	1.6700801
MDLAGMIP	15.681057	10.734128
MDRASKEY	260.32406	26.598875
MDRASBIG	275.16736	0.996981
MDRASMIP	294.57246	10.622034
RESMIN	333.99010	12.231047

5301.84 and 5325.56 for one class scenario. The appropriate χ^2's for direct and inverse coefficients are calculated from the results of regression equations and presented in tables 8-7 and 8-8. The distribution of χ^2's are presented in tables C-21 to C-40 of Appendix (C) and graphically depicted in figures F-20 to F-40 of Appendix (F).

In the case of forty classes direct coefficients and at

TABLE 8-8

CHI SQUARE TEST, INDIVIDUAL COEFFICIENTS

	DIRECT COEFFICIENTS	INVERSE COEFFICIENTS
MDNAVKEY	76.41038	17.46143
MDNAVBIG	74.49140	13.14949
MDNAVMIP	74.66268	11.87695
MDLAGKEY	54.30283	10.79940
MDLAGBIG	20.65279	6.00445
MDLAGMIP	23.64511	7.31354
MDRASKEY	52.26998	10.61322
MDRASBIG	20.33984	6.12268
MDRASMIP	28.39999	7.28215
RESMIN	72.95339	12.30575

1% level, the null hypothesis can be rejected for all estimates except those of MDLAGMIP and MDLAGBIG. At 0.5% level, the MDLAGKEY will be included in this list. The implication being absence of similarity among the estimated matrices and the actual matrix, except in the cases mentioned. The ranking, accordingly, will be MDLAGBIG, MDLAGMIP, MDLAGKEY, MDNAVKEY, MDRASKEY, MDRASBIG, MDNAVBIG, MDNAVMIP, MDRASMIP, and RESMIN.

For inverse matrices, at both levels of significance, one fails to reject the null hypothesis in all instances, which indicates existence of similarity between all estimated inverse matrices and the actual one. The ranking in this case is MDRASBIG, MDLAGBIG, MDNAVBIG, MDRASMIP, MDLAGMIP, RESMIN, MDNAVMIP, MDLAGKEY, MDRASKEY, and MDNAVKEY.

In the case of one class data, and at both levels of significance, one will fail to reject the null hypothesis for all direct and inverse estimates. This, in turn, may be taken as an indication of existence of similarity among estimates and actual data. The ranking in this case is MDRASBIG, MDLAGBIG, MDLAGMIP, MDRASMIP, MDRASKEY, MDLAGKEY, RESMIN, MDNAVBIG, MDNAVMIP, and MDNAVKEY for the direct coefficients and MDLAGBIG, MDRASBIG, MDLAGMIP, MDRASMIP, MDRASKEY, MDLAGKEY, MDNAVMIP, RESMIN, MDNAVBIG, and MDNAVKEY in the case of inverse coefficients.

The results obtained here are generally in agreement with those obtained in unmodified cases. In the one class experiment, all methods for direct and inverse coefficients, similar to the unmodified scheme, showed estimates that are similar to the actual table. In the forty classes experiment the inverse cases, like their counterparts under the unmodified scenario, demonstrate closer estimation of the actual table than did the direct ones, and they all passed the chi-square test. In the case of direct coefficients, and under forty classes scenario, in both modified and unmodified versions, only FRIED and its variants provided estimates similar to the actual table.

It worth noting that all modified estimates, even for the cases that the null hypothesis could be rejected, the χ^2 values declined as compared to the unmodified cases. This, in itself, was no surprise since the modified cases contain 497 actual coefficients and, *ceteris paribus*, must be more similar to the actual matrix. However, as was discussed earlier, inclusion of exogenous data may cause deterioration of estimates for the remaining coefficients. Now, what was interesting here is the fact that this expectation did not materialize. That is, inclusion of exogenous data, did not caused estimates of the remaining coefficients to deteriorate. All other points and explanatory notes mentioned in chapter six with regards to chi-square test are relevant here as well, hence will not be repeated.

8.2.6) Mean Absolute Deviation. The first column of table 8-9 depict the MAD statistic for direct coefficients of estimations obtained via modified methods. Although these values in themselves can not be accurately judged, one can still infer that they are generally low. The values of MAD statistic range from low of 0.00076 for MDRASBIG to high of 0.00203 for RESMIN. The ranking based on MAD for direct coefficients is MDRASBIG, MDNAVBIG, MDLAGBIG, MDRASMIP,

TABLE 8-9

ABSOLUTE AND RELATIVE MEASURES OF FORECASTING ACCURACY

ACTUAL AND PREDICTED DIRECT COEFFICIENTS

	MAD	STPE	RMS	U	UM	US	UC
MDNAVKEY	0.00194	0.25673	0.06825	0.11237	0.00001	0.00032	0.99967
MDNAVBIG	0.00077	0.10155	0.02241	0.03756	0.00024	0.00000	0.99975
MDNAVMIP	0.00115	0.15261	0.03792	0.06401	0.00009	0.00021	0.99970
MDLAGKEY	0.00181	0.23933	0.05735	0.09466	0.00000	0.00030	0.99970
MDLAGBIG	0.00078	0.10299	0.02164	0.03623	0.00000	0.00000	1.00000
MDLAGMIP	0.00108	0.14378	0.03768	0.06291	0.00000	0.00002	0.99998
MDRASKEY	0.00175	0.23249	0.05523	0.09135	0.00000	0.00024	0.99976
MDRASBIG	0.00076	0.10118	0.02141	0.03585	0.00000	0.00000	1.00000
MDRASMIP	0.00106	0.14057	0.03650	0.06094	0.00000	0.00002	0.99998
RESMIN	0.00203	0.26969	0.06185	0.10179	0.00000	0.00039	0.99961

MDLAGMIP, MDNAVMIP, MDRASKEY, MDLAGKEY, MDNAVKEY, and RESMIN.

The MAD values for estimated inverse coefficients, as is evident from table 8-10, basically follow the same pattern. They range from low of 0.00130 for MDRASBIG to high of 0.00384 in the case of MDNAVKEY.

The ranking of all methods for inverse coefficients, as measured by MAD, is MDRASBIG, MDLAGBIG, MDRASMIP, MDNAVBIG, MDLAGMIP, MDNAVMIP, MDRASKEY, MDLAGKEY, RESMIN, and MDNAVKEY.

TABLE 8-10

ABSOLUTE AND RELATIVE MEASURES OF FORECASTING ACCURACY

ACTUAL AND PREDICTED INVERSE COEFFICIENTS

	MAD	STPE	RMS	U	UM	US	UC
MDNAVKEY	0.00384	0.12693	0.11469	0.04661	0.00007	0.00012	0.99981
MDNAVBIG	0.00177	0.05833	0.03687	0.01494	0.00106	0.00009	0.99885
MDNAVMIP	0.00255	0.08442	0.06624	0.02703	0.00058	0.00119	0.99824
MDLAGKEY	0.00291	0.09623	0.08486	0.03426	0.00000	0.00004	0.99996
MDLAGBIG	0.00131	0.04334	0.03013	0.01219	0.00000	0.00001	0.99999
MDLAGMIP	0.00179	0.05906	0.05463	0.02207	0.00000	0.00003	0.99997
MDRASKEY	0.00284	0.09398	0.08179	0.03301	0.00000	0.00006	0.99994
MDRASBIG	0.00130	0.04295	0.03003	0.01215	0.00000	0.00001	0.99999
MDRASMIP	0.00175	0.05777	0.05235	0.02115	0.00000	0.00006	0.99994
RESMIN	0.00338	0.11162	0.09497	0.03826	0.00000	0.00016	0.99984

8.2.7) **Standardized Total Percentage Error**. The second columns of tables 8-9 and 8-10 represent STPE or the average of absolute differences between estimated and actual coefficients as percentage of the mean of the actual matrix, for direct and inverse coefficients. Just like the MAD statistic, the values of STPE for all estimates are relatively low in both direct and inverse estimates. These values, for direct coefficients, range from 0.10118 for MDRASBIG to 0.26969 for RESMIN and yield a ranking that is identical to the ranking by MAD statistic, i.e., MDRASBIG,

MDNAVBIG, MDLAGBIG, MDRASMIP, MDLAGMIP, MDNAVMIP, MDRASKEY, MDLAGKEY, MDNAVKEY, and RESMIN.

Through examination of table 8-10, it will be apparent that STPE values for inverse coefficients range from 0.04295 in the case of MDRASBIG to 0.12693 for MDNAVKEY. The ranking is MDRASBIG, MDLAGBIG, MDRASMIP, MDNAVBIG, MDLAGMIP, MDNAVMIP, MDRASKEY, MDLAGKEY, RESMIN, and MDNAVKEY, which is identical to that of MAD for inverse coefficients.

8.2.8) **Root Mean Square**. This statistic for direct and inverse coefficients is reported in columns three of tables 8-9 and 8-10. They are all relatively low and acceptable. The lowest RMS values are obtained in the case of MDRASBIG. These values for direct and inverse coefficients are 0.02141 and 0.03003 respectively. The largest RMS values are those of MDNAVKEY, i.e., 0.06825 for direct and 0.11469 for inverse coefficients.

The ranking of estimation techniques, in accordance with RMS values, for both direct and inverse tables is the same. Namely, MDRASBIG, MDLAGBIG, MDNAVBIG, MDRASMIP, MDLAGMIP, MDNAVMIP, MDRASKEY, MDLAGKEY, RESMIN and MDNAVKEY.

8.2.9) **Theil's U**. The calculated U statistic, in the case of direct coefficients, for estimation methods utilized here assumes values ranging from 0.03585 associated with MDRASBIG to 0.11237 generated by MDNAVKEY. The same values for inverse coefficients are 0.01215 calculated for MDRASBIG to o.04661 obtained for MDNAVKEY. Hence, it may be inferred that the U values are low enough and acceptable.

The forth columns of tables 8-9 and 8-10 portrays the U values for estimated direct and inverse matrices. Both instances have identical ranking of MDRASBIG, MDLAGBIG, MDNAVBIG, MDRASMIP, MDLAGMIP, MDNAVMIP, MDRASKEY, MDLAGKEY, RESMIN and MDNAVKEY.

8.2.10) UM. Measure of systematic error in the estimates, i.e., UM, is tabulated in the fifth columns of tables 8-9 and 8-10. In both cases of direct and inverse coefficients, the UM values are zeros for all methods except the three variants of the NAIVE method. The numerical values of UM for these latter three instances range from 0.00001 to 0.00024 for direct coefficients and 0.00007 to 0.00106 for inverse coefficients. The ranking of estimation methods based on UM values is MDRASBIG & MDLAGBIG & MDRASMIP & MDLAGMIP & MDRASKEY & MDLAGKEY & RESMIN, MDNAVKEY, MDNAVMIP, and MDNAVBIG.

8.2.11) US. The covariance proportions of calculated Theil's U for the estimation techniques are represented in the sixth columns of tables 8-9 and 8-10. For direct coefficients, the US values oscillate between zero and 0.00039. The range for inverse coefficients is 0.00001 and

0.00039. The resultant rankings are MDRASBIG & MDLAGBIG & MDNAVBIG, MDRASMIP & MDLAGMIP, MDNAVMIP, MDRASKEY, MDLAGKEY, MDNAVKEY, and RESMIN for the direct estimated matrices and MDRASBIG & MDLAGBIG, MDLAGMIP, MDLAGKEY, MDRASKEY & MDRASMIP, MDNAVBIG, MDNAVKEY, RESMIN, and MDNAVMIP for the inverse coefficients.

8.2.12) **UC**. The measure of unsystematic error or covariance proportions of the Theil's U for estimated matrices are reported in the seventh columns of tables 8-9 and 8-10. For direct coefficients the UC statistics range from 1 to 0.99961 which yield the ranking of MDRASBIG & MDLAGBIG, MDLAGMIP & MDRASMIP, MDRASKEY, MDNAVBIG, MDNAVMIP & MDLAGKEY, MDNAVKEY, and RESMIN. The range of UC for inverse coefficients is from 0.99999 to 0.99824 which leads to the ranking of MDRASBIG & MDLAGBIG, MDLAGMIP, MDLAGKEY, MDRASMIP & MDRASKEY, RESMIN, MDNAVKEY, MDNAVBIG, and MDNAVMIP.

8.2.13) **Mean of the Estimated Coefficients**. Previous statistics are heavily influenced by presence of large errors, which most likely are due to existence of very small coefficients in the original table. To supplement these statistics and somewhat compensate for the bias, means, standard deviations, and maximum values of the coefficients of estimated matrices are calculated and reported in tables

8-11 and 8-12.

In the case of direct coefficients all methods, except the three versions of NAIVE method, generate matrices with means identical to that of the actual table. The exceptions are MDNAVKEY, MDNAVBIG, and MDNAVMIP which underestimate the actual mean by 2.3%, 4.51%, and 4.64% respectively.

In the case of inverse coefficients the mean of RESMIN is identical to the mean of the actual table. MDLAGBIG, MDRASKEY, and MDLAGKEY each overestimate the actual mean by 0.03% while MDRASBIG overestimates it by 0.07%. MDRASMIP, MDLAGMIP, MDNAVKEY, MDNAVBIG, and MDNAVMIP underestimate the mean by 0.1%, 0.13%, 3.17%, 3.97%, and 4.48% respectively.

8.2.14) <u>Standard Deviation of the Estimated Coefficients</u>. In the case of direct coefficients, MDRASBIG and MDLAGBIG each underestimate the standard deviation of the actual data by 0.12%, while MDNAVBIG underestimates it by 0.14%. MDLAGMIP and MDRASMIP each overestimate the actual standard deviation by 0.52%. Out of the remaining methods, MDNAVMIP underestimate the actual standard deviation by 1.59%, while MDRASKEY, MDLAGKEY, MDNAVKEY, and RESMIN overestimate this value by 2.45%, 2.89%, 3.52%, and 3.52% respectively.

Overestimation of the standard deviation of the actual inverse matrix are 0.20%, 0.29%, 0.36%, 0.46%, and 0.84% by MDLAGMIP, MDRASMIP, MDLAGKEY, MDRASKEY, and RESMIN, while the

TABLE 8-11

MEAN AND STANDARD DEVIATION OF DIRECT COEFFICIENTS

	MEAN	STAN.DEV.	MAXIMUM
ACTUAL	0.00754	0.03465	0.70512
MDNAVKEY	0.00737	0.03587	0.90439
MDNAVBIG	0.00720	0.03460	0.70512
MDNAVMIP	0.00719	0.03410	0.70512
MDLAGKEY	0.00754	0.03565	0.71433
MDLAGBIG	0.00754	0.03461	0.70512
MDLAGMIP	0.00754	0.03483	0.70512
MDRASKEY	0.00754	0.03550	0.71108
MDRASBIG	0.00754	0.03461	0.70512
MDRASMIP	0.00754	0.03483	0.70512
RESMIN	0.00754	0.03587	0.71802

underestimations by MDRASBIG, MDLAGBIG, MDNAVBIG, MDNAVKEY, and MDNAVMIP are 0.08%, 0.08% 0.24%, 0.87%, and 1.59%.

8.3.15) **Maximum Value of the Estimated Coefficients**. The maximum values of coefficients generated by six of the estimation techniques, i.e., MDRASBIG, MDRASMIP, MDLAGBIG, MDLAGMIP, MDNAVBIG, and MDNAVMIP are identical to the maximum value of the actual coefficients. This outcome, of

TABLE 8-12

MEAN AND STANDARD DEVIATION OF INVERSE COEFFICIENTS

	MEAN	STAN.DEV.	MAXIMUM
ACTUAL	0.03026	0.14357	2.05370
MDNAVKEY	0.02930	0.14232	1.84333
MDNAVBIG	0.02906	0.14322	2.05108
MDNAVMIP	0.02867	0.14129	2.05113
MDLAGKEY	0.03027	0.14408	2.05203
MDLAGBIG	0.03027	0.14345	2.05312
MDLAGMIP	0.03022	0.14386	2.05308
MDRASKEY	0.03027	0.14421	2.04308
MDRASBIG	0.03028	0.14346	2.05296
MDRASMIP	0.03023	0.14398	2.05304
RESMIN	0.03026	0.14477	2.09141

course, was expected since 497 exogenously determined coefficients in these cases were either largest or most important coefficients which were taken from the actual table. The results, however, can be another indication in support of the notion that the largest coefficients in an I-O table are the most important coefficients in updating procedures. The remaining methods, namely MDRASKEY, MDLAGKEY, RESMIN, and MDNAVKEY overestimate the maximum value by 0.82%, 1.31%, 1.83%, and 28.26% respectively.

Maximum values, in the case of inverse coefficients, are underestimated via MDRASMIP. MDLAGMIP, MDLAGBIG, MDRASBIG, MDLAGKEY, MDNAVMIP, MDNAVBIG, MDRASKEY, and MDNAVKEY by 0.03%, 0.03%, 0.03%, 0.04%, 0.08%, 0.13%, 0.13%, 0.52% and 10.24% respectively. RESMIN overestimates the maximum value by 1.84%.

8.3) SUMMARY OF THE RESULTS:

Comparison of tables 6-9 and 8-9 reveals that incorporation of exogenous data improves absolute and relative measures of forecasting accuracy for direct coefficients. All statistics drastically improve for all modified versions. This is particularly true for the cases when the largest 497 coefficients were exogenously obtained and incorporated in the tables. This drastic improvement is yet another indication of the heavy influence of relatively small number of large coefficients on the whole table, since an accurate estimation of these coefficients substantially improved the final estimated tables. It is interesting to note that in some cases (i.e., MDNAVKEY, MDNAVMIP, MDRASKEY, and MDLAGKEY) although measures of closeness improves with modification, but some deterioration occurs in the US and UC statistics. This result, again, is in accordance with previous findings. That is, in some instances, incorporation of exogenous data, while improving the overall accuracy of tables, may cause worsening of the estimates for the remaining coefficients. The worsening will be particularly relevant for very small coefficients. The exact same results for inverse coefficients may be obtained via comparison of tables 6-10 and 8-10, with same explanation and consequences as in the case of direct coefficients.

It should be noted that, in both cases of inverse and direct coefficients, rather drastic improvements are gained when the "large coefficient" and "most important" criterions are utilized to identify the exogenously determined cells. The improvement, however, is much more modest in the cases when the "key coefficient" criterion is employed. This observation, once more, emphasizes the crucial role of a limited number of coefficients in the whole table and reiterates the fact that researchers are best advised to apply their limited resources to identify and exogenously estimate these cells in order to obtain more accurate updated tables.

Comparison of tables 6-11 and 8-11 and 6-12 and 8-12 reveals the effects of exogenous information on the means, the standard deviations, and the maximum values of the updated tables for both direct and inverse coefficients. It is evident from the comparison that the tables in both cases are virtually identical. That is to say, the modification did not have a material effect on the means, the standard deviations, and the maximum values of the estimated tables. This result is not surprising if one remembers the modification process, i.e., the central tendency and dispersement of the set of coefficients remains basically the same whether one utilizes exogenous data or not.

Tabular summary of the rankings of the modified estimation methods as measured by the chosen criterions for both direct and inverse coefficients are presented in tables 8-13 and 8-14 below. All notations and explanations are identical to those of tables 6-13 and 6-14 which were explained earlier. All comments and discussion about the results, potential shortcomings, problems, and interpretation of tables 6-13 and 6-14 are pertinent here as well, hence will not be duplicated. Keeping in mind all the points raised with regards to interpretation of the results in chapter 6, one can conclude the followings.

The Lagrangian type methods suffer from the deficiency that they can not guarantee generation of non-negative coefficients. This fact alone, considerably reduces their usefulness. Thus, a researcher must either resort to imposition of non-negativity conditions in the formulation of the problem, which in turn transforms the problem into a quadratic programming one, or determine the number of generated negative coefficients and make an assessment of the extent of adverse effects of these negative coefficients on his/her overall results.

TABLE 8-13

SUMMARY RANKINGS OF MODIFIED ESTIMATION METHODS

DIRECT COEFFICIENTS

METHOD	N.N	W.U	θ/5	θ/20	μ_θ	σ_θ	MAX_θ	F
MDNAVKEY	1	5	8	3	3	3	4	10
MDNAVBIG	1	8	7	8	9	10	9	3
MDNAVMIP	1	9	4	6	5	9	9	6
MDLAGKEY	8	6	9	2	1	2	2	8
MDLAGBIG	10	1	5	9	8	7	5	2
MDLAGMIP	9	6	2	5	6	5	7	5
MDRASKEY	1	3	6	1	2	1	1	7
MDRASBIG	1	1	2	7	10	8	6	1
MDRASMIP	1	3	1	4	7	6	8	4
RESMIN	1	10	10	10	4	4	3	9

TABLE 8-13 CONTINUED

SUMMARY RANKINGS OF MODIFIED ESTIMATION METHODS

DIRECT COEFFICIENTS

METHOD	t_α	t_β	R^2	$\chi^{2/1}$	$\chi^{2/40}$	J.F	MAD	STP
MDNAVKEY	8	4	10	10	4	4	9	9
MDNAVBIG	10	8	3	8	7	9	2	2
MDNAVMIP	9	10	6	9	8	10	6	6
MDLAGKEY	4	5	8	6	3	2	8	8
MDLAGBIG	6	6	1	2	1	7	3	3
MDLAGMIP	2	2	5	3	2	1	5	5
MDRASKEY	3	3	7	5	5	5	7	7
MDRASBIG	5	7	1	1	6	8	1	1
MDRASMIP	1	1	4	4	9	6	4	4
RESMIN	7	9	9	7	10	3	10	10

TABLE 8-13 CONCLUDED

SUMMARY RANKINGS OF MODIFIED ESTIMATION METHODS

DIRECT COEFFICIENTS

METHOD	RMS	U	UM	US	UC	μ_α	MAX_a	σ_a
MDNAVKEY	10	10	8	9	9	8	10	9
MDNAVBIG	3	3	10	1	6	9	1	3
MDNAVMIP	6	6	9	6	7	10	1	6
MDLAGKEY	8	8	1	8	7	1	8	8
MDLAGBIG	2	2	1	1	1	1	1	1
MDLAGMIP	5	5	1	4	3	1	1	4
MDRASKEY	7	7	1	7	5	1	7	7
MDRASBIG	1	1	1	1	1	1	1	1
MDRASMIP	4	4	1	4	3	1	1	4
RESMIN	9	9	1	10	10	1	9	9

TABLE 8-14

SUMMARY RANKINGS OF MODIFIED ESTIMATION METHODS

INVERSE COEFFICIENTS

METHOD	N.N	W.U	θ/5	θ/20	μ_θ	σ_θ	MAX_θ	F
MDNAVKEY	1	10	9	10	4	10	10	10
MDNAVBIG	1	8	7	8	7	9	9	3
MDNAVMIP	1	8	8	7	10	1	1	6
MDLAGKEY	1	4	6	6	6	8	8	8
MDLAGBIG	10	1	4	4	8	4	4	2
MDLAGMIP	1	4	2	2	2	3	3	5
MDRASKEY	1	4	5	5	5	7	7	7
MDRASBIG	1	1	3	3	9	5	5	1
MDRASMIP	1	1	1	1	1	2	2	4
RESMIN	1	4	10	9	3	6	6	9

TABLE 8-14 CONTINUED

SUMMARY RANKINGS OF MODIFIED ESTIMATION METHODS

INVERSE COEFFICIENTS

METHOD	t_α	t_β	R^2	$\chi^{2/1}$	$\chi^{2/40}$	J.F	MAD	STP
MDNAVKEY	8	9	10	10	10	8	10	10
MDNAVBIG	10	8	1	9	3	9	4	4
MDNAVMIP	9	10	6	7	7	10	6	6
MDLAGKEY	1	1	7	6	8	1	8	8
MDLAGBIG	3	5	1	1	2	7	2	2
MDLAGMIP	3	2	4	3	5	2	5	5
MDRASKEY	2	3	7	5	9	4	7	7
MDRASBIG	5	4	1	2	1	3	1	1
MDRASMIP	7	6	4	4	4	5	3	3
RESMIN	6	7	9	8	6	6	9	9

TABLE 8-14 CONCLUDED

SUMMARY RANKINGS OF MODIFIED ESTIMATION METHODS

INVERSE COEFFICIENTS

METHOD	RMS	U	UM	US	UC	μ_α	MAX_a	σ_a
MDNAVKEY	10	10	8	8	8	8	10	9
MDNAVBIG	3	3	10	7	9	9	7	4
MDNAVMIP	6	6	9	10	10	10	7	10
MDLAGKEY	8	8	1	4	4	2	5	6
MDLAGBIG	2	2	1	1	1	2	1	1
MDLAGMIP	5	5	1	3	3	7	1	3
MDRASKEY	7	7	1	5	5	2	6	7
MDRASBIG	1	1	1	1	1	5	4	1
MDRASMIP	4	4	1	5	5	6	1	5
RESMIN	9	9	1	9	7	1	9	8

In the context of the current project, three methods, namely MDLAGKEY, MDLAGBIG, and MDLAGMIP, did generate negative coefficients. The number of such negative coefficients, however, is very small when compared to the total number of coefficients in the table. Hence, this problem, in the case of current research, does not pose a serious problem.

All of the methods utilized here, show acceptable results based on the Wilcoxon test and no one method stands out as superior to the others based on this test. Tests related to coefficient of equality reveals that the RESMIN method is less powerful than the other methods, hence may be discarded from further consideration. Regression related tests, although not conclusively, point out that while all of the methods are basically performing the same, variants of the NAIVE method suffer some setbacks and "MIP" and "BIG" variants of the RAS and the LAGRANGIAN perform better than the rest. The χ^2 tests also yield the same results with some advantage given to the LAGRANGIAN type methods.

It is consideration of the remaining closeness criterions that determines the final ranking of the updating techniques. MDNAVKEY, MDLAGKEY, and MDRASKEY, by virtue of their relative performances may be ranked as number nine, eight, and seven respectively and be dropped from further consideration. The overall ranking for the rest of updating methods, although in some cases not clear cut, may be stated as MDRASBIG, MDLAGBIG, MDNAVBIG, MDRASMIP, MDLAGMIP, and MDNAVMIP. In the case of inverse coefficients the ranking is fundamentally the same, except RESMIN and MDNAVKEY exchange their rankings, thereby making the MDNAVKEY the least powerful updating method.

Hence, it seems that relatively good updates can be obtained via some of these updating techniques and incorporation of exogenous data substantially improves the overall estimates. This is particularly true when the "BIG" and "MIP" coefficients are selected as the exogenously determined cells. Among the methods utilized in this study, RAS and LAGRANGIAN methods are the most efficient methods, especially when they are modified via the "BIG" criterion. The same criterion applied to NAIVE method, however, provides very good and compatible results, which can be an indication of relative stability of the I-O coefficients in short and medium runs. Thus, presenting the NAIVE method, when modified via exogenous estimation of some subset of its largest coefficients, as a viable alternative.

As was mentioned earlier, all the points and concerns raised in chapter six with regards to the ranking and evaluation of updating techniques are relevant here as well. Nevertheless, it merits to especially note two points again. First, the closeness tests are used here only as a mean of *ranking* the updating techniques and not as absolute tests of "goodness" or efficiency. Second, as it is evident from the results, these updating methods do not estimate all of the coefficients with the same degree of accuracy. That is to say, a subset of coefficients are estimated rather accurately, while some others are estimated with relatively substantial errors.

CHAPTER NINE

AGGREGATION

"Out with it, Tarrou! What on earth prompted you to take a hand in this?"
"I don't know. My code of morals, perhaps."
"Your code of morals? What code?"
"Comprehension."

Albert Camus

9.1) INTRODUCTION:

One of the primary concerns of input-output analysts is the issue of aggregation. There are many instances, where constraints imposed on researchers, mandates construction or utilization of smaller tables. The question, under these circumstances, is the trade off between loss of accuracy due to aggregation in one hand, and the gain in economizing resources on the other. How large is that loss and what are the gains, is the main concern of aggregation literature. Probing this issue is beyond the scope of the current study, and will not be pursued here. However, only one part of the aggregation question will be briefly explored here, since it properly correlates with the overall purpose of present project and will set the stage for subsequent analysis at a later date.

Theoretically, if one can have a table containing one sector for every commodity or firm in the economy, aggregation and its related problems cease to exist. Practical considerations, however, do not permit this case. Short of this ideal situation, then, and regardless of the size of the table selected, some degree of aggregation will be present in any actual table. This stems from the fact that, in theory, only commodities or firms that utilize identical production functions may be classified in one sector. To the extent that this is not possible, and to the extend that commodities or firms with different production functions are classified in one sector, "impurities" are introduced in the representative production functions of each sector. For this will be only an "average" production function, representing all commodities or firms classified in that particular sector. This is what I-O analysts refer to as "aggregation bias." A vast body of literature is devoted to the treatment of this bias and its relevant issues, which came to be known as the "aggregation problem."

Hatanaka (1952) was probably the first author addressing this issue. Following him, numerous researchers attempted to tackle the problem. OhUallachain (1985), Chakraborty and ten Raa (1981), both cited in Lahr and Stevens (1987), and Kymn (1990) are among those who provide comprehensive reviews of this literature.

In the beginning, the principal problem was to decide which firms or commodities should be classified in one sector. Originally, data collection and computational capabilities were among major reasons for being concerned about the size of the input-output tables. Improvement in statistical and computational capabilities have largely alleviated these problems, and the debate on these issues have subsided substantially. Advances in computing facilities have substantially enhanced the computing abilities of input-output researchers. Moreover, many countries constructed input-output tables for several years and the sectoral structure of these tables are basically established and relatively constant. Thus, a more pressing problem is consistent aggregation and aggregation of already existing tables.

Another issue in input-output aggregation, is the effects of aggregation on the stability of coefficients. For if aggregated tables are more stable, then there may be no need for updating techniques in case of relatively small tables. The *a priori* expectation, however, is inconclusive.

For theoretically, aggregation should attenuate the substitution effects and lead to greater stability of coefficients. Disaggregated tables, on the other hand, should demonstrate greater coefficient stability due to increase in homogeneity of the production processes classified in each sector. Researchers such as Karaska (1968), Doeksen and Little (1968), Tiebout (1969), and Carter (1970) have explored this question with mixed conclusions.

Many researchers have explored various aggregation schemes, conditions, assumptions, and methods that allow performing aggregation with zero bias, aggregation with minimum bias, measurement of the bias, etc. A partial list of such researchers who investigated and addressed aggregation issues includes Yamada (1961), Green (1964), Doeksen and Little (1968), Karaska (1968), Morimoto (1970), Williamson (1970), Hewings (1972), Ijiri (1971), Kossov (1972), Stevens and Trainer (1976), Isserman (1977), Bulmer-Thomas (1982), Gibbons, Wolsky, and Tolley (1982), Sawyer and Miller (1983), Ralston, Hastings, and Brucker (1986), Crown (1987), and (1990), Howe and Johnson (1987, cited in Lahr and Stevens, 1987), Garhart and Giarratani (1987), and Lahr and Stevens (1987).

The results of these investigations, however, are not conclusive and uniform. Although some researchers (e.g., Crown, 1987) believe that aggregation bias primarily is an accounting issue that can be completely eliminated in the base year, a dominant consensus, nevertheless, is evident. That is, the aggregation introduces a bias in the model and the level of aggregation directly affects this bias. Theoretically, it may be possible and desirable to eliminate this bias via construction of highly detailed tables. However, in many cases some level of aggregation is necessitated by various circumstances and thereby the results of the I-O analysis is affected.

Full exploration of the aggregation area is enticing and potentially fruitful, but will not be pursued here. Instead, a limited probe will be carried out with specific aims. Namely, analysis of the effects of aggregation on the stability of coefficients and relative performance of updating methods. This will be achieved via application of selected updating methods to aggregated data.

Aggregation here is achieved through usage of a simple "aggregation matrix." This matrix, denoted as (S), is a (k) by (n) matrix containing ones and zeros, where (k) and (n) represent the number of sectors in the aggregated and original tables respectively. All elements of (S) are zeros except one element in each column that is (1). The place of ones in each row (i) of the (S) matrix designates the sectors of the original table that are to be included in the i-th sector of the aggregated table. Designating Z, Z^*, Y, and Y^* as the original transaction matrix, aggregated transaction matrix, original final demand vector, and aggregated demand vector respectively, one can have:

$$Y^{*} = (S)(Y) \tag{9-1}$$

and

$$Z^{*} = (S)(Z)(S)' \tag{9-2}$$

The new aggregated vector of total output then will be:

$$X^{*} = (Z^{*})(i) + Y^{*} \tag{9-3}$$

where (i) is a column vector of ones.

9.2) IMPLEMENTATION AND RESULTS:

Application of the above procedure to the original tables furnished the required (survey based) aggregated tables. The aggregation was performed at three different levels. The number of sectors at each of these levels are 35, 16, and 6. The "aggregation scheme" for the three levels of aggregation along with the classifications of the "original" table are provided in appendix (G).

Following the aggregation procedure, the three aggregated tables were updated via utilization of the eight updating techniques used in chapter six. Adding a constant coefficient case or NAIVE method to this list, provided nine updated tables for each level of aggregation, which implies total of twenty seven updated tables. Next, inverse coefficients for each table was calculated and then updated tables of both direct and inverse coefficients were compared with the target year's actual aggregated direct and inverse matrices through employment of a selected set of closeness tests. The results are tabulated in tables below.

9.2.1) **Thirty Five Sectors Table**. According to table 9-1, the MAD statistic for the aggregated (35) by (35) updated tables are relatively low and ranges from low of 0.00275 to high of 0.05427. The appropriate ranking of the techniques, then, is RAS, FRIED, RECRAS, NAIVE, PROPVA, RASLAG, RERALA, RECLAG, and ALMON. The same ranking holds for STPE and RMS statistics. They range from 0.18484 to 3.64446 and 0.04561 to 1.40088 respectively. The STPE values also may be considered low.

The ranking based on Theil's U is slightly different as RAS, FRIED, RECRAS, RASLAG, NAIVE, PROPVA, RERALA, RECLAG, and ALMON, with the range of 0.0045 to 0.81671. The same table for UM values indicates range of zero and 0.00036 with the ranking of RAS & FRIED & RECRAS & RASLAG, ALMON, NAIVE, RERALA, RECLAG, and PROPVA. The rankings based on US and UC are identical and may be expressed as FRIED, RAS, NAIVE, PROPVA, RECRAS, RECLAG, RASLAG, RERALA, and ALMON. The respective low high for these two statistics are 0.00007 to 0.01896 and 0.98104 to 0.99993. Except in the case of ALMON,

TABLE 9-1

ABSOLUTE AND RELATIVE MEASURES OF FORECASTING ACCURACY

ACTUAL AND PREDICTED (35) BY (35) DIRECT COEFFICIENTS

	MAD	STPE	RMS	U	UM	US	UC
NAIVE	0.00346	0.23219	0.06849	0.11513	0.00007	0.00058	0.99934
RAS	0.00275	0.18484	0.04561	0.07745	0.00000	0.00008	0.99992
RECRAS	0.00327	0.21970	0.06427	0.10663	0.00000	0.00194	0.99806
PROPVA	0.00358	0.24012	0.06955	0.12182	0.00036	0.00116	0.99848
FRIED	0.00277	0.18634	0.05643	0.07886	0.00000	0.00007	0.99993
RECLAG	0.00453	0.30394	0.08614	0.13811	0.00027	0.00646	0.99327
RASLAG	0.00407	0.27319	0.06967	0.11224	0.00000	0.00748	0.99252
RERALA	0.00423	0.28436	0.08314	0.13252	0.00016	0.00804	0.99180
ALMON	0.05427	3.64446	1.40088	0.81671	0.00001	0.01896	0.98104

all values of U, UM, US, and UC appear to be reasonable.

The means, standard deviations, and the maximum values of the aggregated (35) by (35) updated tables are given in the table 9-2 above. The same statistics for the actual 35 sectors table are included in this table to facilitate the comparison. It is evident that RAS, FRIED, and RASLAG duplicate the mean of the actual table. RECRAS overestimates the mean by 1%. NAIVE,RERALA, PROPVA, RECLAG, and ALMON underestimate the actual mean by 4%, 7%, 9%, 9%, and 22% respectively. (All percentages in this chapter are rounded

TABLE 9-2

MEAN, STANDARD DEVIATION, AND MAXIMUM VALUE
(35) BY (35) PREDICTED DIRECT COEFFICIENTS

ESTIMATION METHOD	MEAN	STANDARD DEVIATION	MAXIMUM VALUE
ACTUAL	0.01489	0.04731	0.59700
NAIVE	0.01430	0.04896	0.70633
RAS	0.01489	0.04772	0.59465
RECRAS	0.01500	0.05013	0.66496
PROPVA	0.01358	0.04494	0.59509
FRIED	0.01489	0.04770	0.59601
RECLAG	0.01348	0.05423	0.71637
RASLAG	0.01489	0.05333	0.61788
RERALA	0.01386	0.05476	0.67849
ALMON	0.01165	0.24017	3.63750

to the nearest whole number). Examination of the standard deviations of the estimated coefficients shows overestimation of 1%, 1%, 3%, 6%, 13%, 15%, 16%, and 408% by RAS, FRIED, NAIVE, RECRAS, RASLAG, RECLAG, RERALA, and ALMON. PROPVA underestimated this statistic by 5%. Thus, the ranking of the methods will be FRIED, RAS, NAIVE, PROPVA, RECRAS, RASLAG, RECLAG, RERALA, and ALMON.

The maximum value of the actual coefficients were matched by FRIED, RAS, and PROPVA. Other methods, namely RASLAG, RECRAS, RERALA, NAIVE, RECLAG, and ALMON, overestimated this value by 3%, 11%, 14%, 18%, 20%, and 509%. The ranking of the updating techniques based on the maximum values of predicted coefficients, therefore, is FRIED, RAS, PROPVA, RASLAG, RECRAS, RERALA, NAIVE, RECLAG, and ALMON.

Tables 9-3 and 9-4 show the same set of statistics for the case of inverse coefficients. Values of all of these statistics seem to be reasonable except in some instances of ALMON method. According to these tables, the MAD and STPE statistics range from 0.00435 to 0.03295 and 0.07152 to 0.54198 respectively, leading to an identical ranking of RAS, FRIED, RECRAS, NAIVE, RASLAG, RERALA, PROPVA, RECLAG, and ALMON.

The RMS statistic ranges from 0.07114 to 0.61763, yielding the ranking of RAS, FRIED, NAIVE, RECRAS, RASLAG, PROPVA, RERALA, RECLAG, and ALMON. The Theil's U with the range of 0.0772 to 0.24349 has the same ranking except that RECRAS and NAIVE switch places. UM ranks the methods as RAS, FRIED, RASLAG, RERALA, RECRAS, RECLAG, ALMON, NAIVE, and PROPVA. The UM ranges from low of zero to the high of 0.00143. The US and UC identically rank the updating techniques as RECRAS, ALMON, RAS, FRIED, NAIVE, PROPVA, RERALA, RASLAG, and RECLAG. The relevant ranges for these

TABLE 9-3

ABSOLUTE AND RELATIVE MEASURES OF FORECASTING ACCURACY
ACTUAL AND PREDICTED (35) BY (35) INVERSE COEFFICIENTS

	MAD	STPE	RMS	U	UM	US	UC
NAIVE	0.00731	0.12018	0.12395	0.04869	0.00055	0.00097	0.99848
RAS	0.00435	0.07152	0.07114	0.02772	0.00000	0.00030	0.99970
RECRAS	0.00694	0.11421	0.12398	0.04809	0.00006	0.00003	0.99991
PROPVA	0.00912	0.14996	0.15518	0.06165	0.00143	0.00271	0.99586
FRIED	0.00440	0.07241	0.07362	0.02870	0.00000	0.00035	0.99965
RECLAG	0.00957	0.15741	0.18687	0.07070	0.00008	0.00472	0.99520
RASLAG	0.00767	0.12607	0.12805	0.04882	0.00002	0.00457	0.99541
RERALA	0.00888	0.14604	0.17456	0.06619	0.00006	0.00449	0.99545
ALMON	0.03295	0.54198	0.61763	0.24349	0.00037	0.00003	0.99960

statistics are 0.00003 to 0.00472 for RECRAS and RECLAG and 0.99520 to 0.99991 for RECLAG and RECRAS respectively.

RAS and FRIED produce means that are identical to the mean of the actual table. RASLAG and RECRAS overestimate the actual mean by 1% and 2% respectively. RERALA, RECLAG, NAIVE, PROPVA, and ALMON underestimate the mean of the actual data by 2%, 3%, 5%, 10%, and 20%. The ranking based on the mean of the coefficients is RAS, FRIED, RASLAG, RERALA, RECRAS, RECLAG, NAIVE, PROPVA, and ALMON.

With regards to the standard deviations of estimated

TABLE 9-4

MEAN, STANDARD DEVIATION, AND MAXIMUM VALUE
(35) BY (35) PREDICTED INVERSE COEFFICIENTS

ESTIMATION METHOD	MEAN	STANDARD DEVIATION	MAXIMUM VALUE
ACTUAL	0.06080	0.20885	2.03280
NAIVE	0.05788	0.20498	1.84298
RAS	0.06082	0.20762	2.04829
RECRAS	0.06172	0.20947	1.77652
PROPVA	0.05493	0.20077	1.91599
FRIED	0.06082	0.20747	2.05237
RECLAG	0.05911	0.22169	2.14986
RASLAG	0.06134	0.21751	2.17349
RERALA	0.05946	0.22055	2.20337
ALMON	0.04886	0.20568	1.44373

coefficients, it can be seen that RECRAS has the same standard deviation as the actual table. RASLAG, RERALA, and RECLAG overestimate this statistic by 4%, 6%, and 6%, while RAS, FRIED, NAIVE, ALMON, and PROPVA underestimate the actual standard deviation by 1%, 1%, 2%, 2%, and 4%.

As far as the maximum values are concerned, it is clear that RAS, FRIED, RECLAG, RASLAG, and RERALA overestimated this value by 1%, 1%, 6%, 7%, and 8%, while the underestimation by PROPVA, NAIVE, RECRAS, and ALMON is 6%, 9%, 13%, and 29%.

Thus, one can conclude that while none of the methods is very accurate on all the coefficients, RAS and FRIED produce estimated direct coefficient tables that, with the exception of some cells, are very close to the original data. RECRAS and NAIVE may be considered as next in line for estimation performance. The latter two methods, as compared to previous two techniques, have larger errors in estimation of certain subset of the coefficients. PROPVA takes the fifth place, with RASLAG and RERALA taking the sixth and seventh places. RECLAG and ALMON may be considered as the eighth and ninth choices. In the inverse coefficient case the overall ranking is RAS, FRIED, RECRAS, NAIVE, RASLAG, RERALA, PROPVA, RECLAG, and ALMON.

9.2.2) **Sixteen Sectors Table**. In the next step, the data was further aggregated into a sixteen sector economy. The results are provided below.

Table 9-5, reveals that for the direct coefficients the MAD and STPE statistics range from low of 0.00446 and 0.13369 associated with RAS to high of 0.01885 and 0.56479 for ALMON, leading to the ranking of RAS, FRIED, RECRAS, NAIVE, PROPVA, RASLAG, RECLAG, RERALA, and ALMON.

RMS with the range of 0.03672 to 0.15448 yields the

TABLE 9-5

ABSOLUTE AND RELATIVE MEASURES OF FORECASTING ACCURACY

ACTUAL AND PREDICTED (16) BY (16) DIRECT COEFFICIENTS

	MAD	STPE	RMS	U	UM	US	UC
NAIVE	0.00558	0.16730	0.04848	0.07212	0.00036	0.00006	0.99958
RAS	0.00446	0.13369	0.03672	0.05457	0.00000	0.00001	0.99999
RECRAS	0.00543	0.16276	0.05032	0.07579	0.00065	0.00118	0.99817
PROPVA	0.00664	0.19904	0.07038	0.10865	0.00176	0.00598	0.99226
FRIED	0.00448	0.13419	0.03788	0.05629	0.00000	0.00002	0.99998
RECLAG	0.00829	0.24829	0.07102	0.09957	0.00005	0.02510	0.97485
RASLAG	0.00742	0.22227	0.06207	0.08836	0.00000	0.01727	0.98273
RERALA	0.00846	0.25344	0.07185	0.10084	0.00009	0.02395	0.97596
ALMON	0.01885	0.56479	0.15448	0.27773	0.00227	0.03456	0.96317

ranking of RAS, FRIED, NAIVE, RECRAS, RASLAG, PROPVA, RECLAG, RERALA, and ALMON.

Theil's U goes from low of 0.0457 to high of 0.27773 and gives the ranking of RAS, FRIED, NAIVE, RECRAS, RASLAG, RECLAG, RERALA, PROPVA, and ALMON. The ranking based on the UM will be RAS & FRIED & RASLAG, RECLAG, RERALA, NAIVE, RECRAS, PROPVA, and ALMON. The US and UC identically rank the estimation methods as RAS, FRIED, NAIVE, RECRAS, PROPVA, RASLAG, RERALA, RECLAG, and ALMON.

Consideration of the means, standard deviations, and

TABLE 9-6

MEAN, STANDARD DEVIATION, AND MAXIMUM VALUE

(16) BY (16) PREDICTED DIRECT COEFFICIENTS

ESTIMATION METHOD	MEAN	STANDARD DEVIATION	MAXIMUM VALUE
ACTUAL	0.03338	0.07728	0.59753
NAIVE	0.03246	0.07765	0.70711
RAS	0.03338	0.07741	0.57556
RECRAS	0.03210	0.07555	0.62159
PROPVA	0.03043	0.07184	0.59246
FRIED	0.03338	0.07746	0.58018
RECLAG	0.03287	0.08853	0.66254
RASLAG	0.03338	0.08544	0.60205
RERALA	0.03270	0.08840	0.66561
ALMON	0.02602	0.04856	0.36128

maximum values of various estimates will shed some more lights on the results.

According to these statistics, as presented in table 9-6, RAS, FRIED, and RASLAG produced the same mean as the actual table. RECLAG and RERALA each underestimated the actual mean by 2%, while NAIVE, RECRAS, PROPVA, and ALMON underestimated the mean by 3%, 4%, 9%, and 22% respectively.

The standard deviations of the estimated tables via RAS, FRIED, and NAIVE matched the standard deviation of the actual table. RECRAS, PROPVA, and ALMON had standard deviations below that of the actual table by 2%, 7%, and 37%, while RASLAG, RERALA, and RECLAG overestimated the actual standard deviation by 11%, 14%, and 15% respectively.

The maximum values of the estimated coefficients by PROPVA, FRIED, RAS, and ALMON were below the actual maximum value by 1%, 3%, 4%, and 40%. The maximum value was overestimated via RASLAG, RECRAS, RECLAG, RERALA, and NAIVE by 1%, 4%, 11%, 11%, and 18% respectively.

Table 9-7 below provides the same statistics in the case of estimated inverse coefficients. The MAD and STPE, while range from 0.00794 to 0.03852 and 0.05844 to 0.28338, identically rank the estimation techniques as RAS, FRIED, RECRAS, NAIVE, RASLAG, PROPVA, RECLAG, RERALA, and ALMON.

The low-high values of RMS and U statistics for the estimated inverse coefficients are 0.07192 and 0.02490 to 0.33332 and 0.12674. Both of these statistics yield ranking of the methods as RAS, FRIED, NAIVE, RASLAG, RECRAS, PROPVA, RERALA, RECLAG, and ALMON. The same table shows the ranking based on the mean of the U statistic to be RERALA, FRIED & RAS & RECLAG, RASLAG, RECRAS, NAIVE, PROPVA, and ALMON. The standard deviations and covariances of U values, provide rankings of FRIED, RAS, RECRAS, NAIVE, PROPVA, RECLAG, ALMON, RERALA, RASLAG, in the case of UM, and FRIED, RAS,

TABLE 9-7

ABSOLUTE AND RELATIVE MEASURES OF FORECASTING ACCURACY
ACTUAL AND PREDICTED (16) BY (16) INVERSE COEFFICIENTS

	MAD	STPE	RMS	U	UM	US	UC
NAIVE	0.01412	0.10388	0.10791	0.03796	0.00351	0.00448	0.99201
RAS	0.00794	0.05844	0.07192	0.02490	0.00003	0.00105	0.99892
RECRAS	0.01375	0.10118	0.14199	0.04986	0.00183	0.00180	0.99637
PROPVA	0.01876	0.13801	0.16553	0.05887	0.00600	0.00629	0.98771
FRIED	0.00803	0.05910	0.07502	0.02597	0.00003	0.00095	0.99902
RECLAG	0.01921	0.14133	0.18450	0.06158	0.00003	0.02835	0.97162
RASLAG	0.01565	0.11515	0.13871	0.04668	0.00011	0.03068	0.96921
RERALA	0.01940	0.14271	0.18105	0.06047	0.00000	0.02885	0.97115
ALMON	0.03852	0.28338	0.33332	0.12674	0.00728	0.02803	0.96469

RECRAS, NAIVE, PROPVA, RECLAG, RERALA, RASLAG, ALMON in the case of UC.

Examination of table 9-8 shows that estimated inverse tables generated via RAS, FRIED, and RERALA have means that match the mean of the actual inverse coefficients, while NAIVE, RECRAS, PROPVA, ALMON underestimate this mean by 5%, 5%, 9%, 21% and RECLAG, and RASLAG overestimate it by 1% respectively.

The same table reveals that RAS, FRIED, RASLAG, RECLAG, and RERALA overestimate the actual standard deviation by 1%,

TABLE 9-8

MEAN, STANDARD DEVIATION, AND MAXIMUM VALUE

(16) BY (16) PREDICTED INVERSE COEFFICIENTS

ESTIMATION METHOD	MEAN	STANDARD DEVIATION	MAXIMUM VALUE
ACTUAL	0.13594	0.33389	2.03620
NAIVE	0.12955	0.32666	1.84590
RAS	0.13635	0.33622	2.05405
RECRAS	0.12987	0.32786	1.76995
PROPVA	0.12312	0.32076	1.91830
FRIED	0.13638	0.33620	2.05850
RECLAG	0.13700	0.36495	2.16099
RASLAG	0.13738	0.35818	2.20529
RERALA	0.13631	0.36464	2.23642
ALMON	0.10750	0.27808	1.44171

1%, 7%, 9%, and 9% while NAIVE, RECRAS, PROPVA, and ALMON generate standard deviations that are below the actual standard deviation by 2%, 2%, 4%, and 17%. Based on the same table, the maximum values of the estimated coefficients are lass than the actual maximum value in case of PROPVA, NAIVE, RECRAS, and ALMON by 6%, 9%, 13%, and 29%, while the same value is overestimated by RAS, FRIED, RECLAG, RASLAG, and

RERALA in the amount of 1%, 1%, 6%, 8%, and 10%.

Overall, for direct coefficients, it may be said that RAS and FRIED take the first and second places in the ranking of the estimation methods. With some distance, RECRAS and NAIVE can be considered as third and forth methods. The fifth and sixth places goes to PROPVA and RASLAG. The seventh, eighth, and ninth places go to RECLAG, RERALA, and ALMON.

In the case of inverse coefficients the ranking is exactly the same, except PROPVA and RASLAG switch places. The values of all closeness criterions for both direct and inverse cases, with the exception of some instances for ALMON method, appear to be reasonable and acceptable.

9.2.3) **Six Sectors Table**. The last stage is examination of the updated results obtained from the aggregated (6) by (6) table. Table 9-9 provides the results of selected accuracy measures. It is evident from the table that MAD statistic ranges from 0.00521 to 0.16161 for RAS and ALMON respectively, yielding the ranking of the estimation techniques as RAS, FRIED, RASLAG, NAIVE, RECLAG, PROPVA, RERALA, RECRAS, and ALMON. The ranking based on STPE is identical to the above ranking with low of 0.08406 to high of 2.60286 associated with RAS and ALMON.

RMS and Theil's U statistics also rank the methods the same as above, with the exception that in

TABLE 9-9

ABSOLUTE AND RELATIVE MEASURES OF FORECASTING ACCURACY

ACTUAL AND PREDICTED (6) BY (6) DIRECT COEFFICIENTS

	MAD	STPE	RMS	U	UM	US	UC
NAIVE	0.00818	0.13211	0.03576	0.05359	0.01058	0.05375	0.93567
RAS	0.00521	0.08406	0.01924	0.02782	0.00000	0.00350	0.99650
RECRAS	0.01067	0.17233	0.05874	0.09096	0.02019	0.07263	0.90718
PROPVA	0.01006	0.16251	0.04684	0.07124	0.01664	0.06144	0.92192
FRIED	0.00524	0.08462	0.01941	0.02804	0.00000	0.00470	0.99530
RECLAG	0.00913	0.14749	0.03906	0.05620	0.01138	0.01437	0.97425
RASLAG	0.00815	0.13160	0.03116	0.04439	0.00000	0.03595	0.96405
RERALA	0.01065	0.17200	0.04325	0.06219	0.01494	0.01509	0.96997
ALMON	0.16121	2.60286	0.82089	0.66719	0.00023	0.08210	0.91767

these latter cases PROPVA and RERALA exchange places in the rankings. The low-high for RMS and U are from low of 0.01924 and 0.02782 to high of 0.82089 and 0.66719.

The UM statistic ranges from zero to 0.02019 with the ranking of RAS, FRIED, RASLAG, ALMON, NAIVE, RECLAG, RERALA, PROPVA, and RECRAS. The US deviation has the low-high of 0.00350 to 0.91767 with the ranking of RAS, FRIED, RECLAG, RERALA, RASLAG, NAIVE, PROPVA, RECRAS, and ALMON. The UC has the low-high of 0.90718 to 0.99650 and yields the same ranking as US, with the exception that RECRAS and ALMON

TABLE 9-10

MEAN, STANDARD DEVIATION, AND MAXIMUM VALUE

(6) BY (6) PREDICTED DIRECT COEFFICIENTS

ESTIMATION METHOD	MEAN	STANDARD DEVIATION	MAXIMUM VALUE
ACTUAL	0.06193	0.12810	0.50337
NAIVE	0.05826	0.11981	0.46645
RAS	0.06193	0.12924	0.51207
RECRAS	0.05359	0.11227	0.46188
PROPVA	0.05589	0.11649	0.47926
FRIED	0.06193	0.12943	0.51237
RECLAG	0.05777	0.13278	0.53087
RASLAG	0.06193	0.13401	0.53542
RERALA	0.05665	0.13341	0.54060
ALMON	0.04941	0.36331	1.71765

exchange places.

The means, standard deviations, and the maximum values of the direct coefficients of estimated tables are presented in table 9-10. It is clear that RAS, FRIED, and RASLAG methods produce means identical to that of the actual table. NAIVE, RECLAG, RERALA, PROPVA, RECRAS, and ALMON underestimate this mean by 6%, 7%, 9%, 10%, 13%, and 20%.

NAIVE, PROPVA, and RECRAS have standard deviations which are 6%, 9%, and 12% less than the actual standard deviation. RAS, FRIED, RECLAG, RERALA, RASLAG, and ALMON have standard deviations that are 1%, 1%, 4%, 4%, 5%, and 184% greater than that of the actual table.

With regards to the maximum values, RAS, FRIED, RECLAG, RASLAG, RERALA, and ALMON exceed the actual maximum value by 2%, 2%, 5%, 6%, 7%, and 241%, while PROPVA, NAIVE, and RECRAS fall short of the actual maximum value by 5%, 7%, and 8% respectively.

Table 9-11 presents the results for the inverse matrices. Accordingly, the MAD statistics range from 0.00888 to 0.13510. The appropriate ranking will be RAS, FRIED, RASLAG, NAIVE, RECLAG, RERALA, PROPVA, RECRAS, and ALMON.

The STPE has the same ranking with range of 0.02894 to 0.44010. The RMS statistic also has the same ranking except that NAIVE and RECLAG switch places. The low-high for RMS is 0.03365 to 0.62763.

The U statistic move from 0.01115 to 0.21960 with the subsequent ranking of RAS, FRIED, RASLAG, RECLAG, RERALA, NAIVE, PROPVA, RECRAS, and ALMON. The means of U statistics range from 0.00001 to 0.03008 with the ranking of RASLAG, FRIED, RAS, ALMON, RECLAG, RERALA, NAIVE, RECRAS, and PROPVA. The rankings for US and UC are ALMON, FRIED, RAS, PROPVA, NAIVE, RASLAG, RECRAS, and RERALA for US, and ALMON, FRIED, RAS, RASLAG, RECLAG, RERALA, PROPVA, NAIVE, and

TABLE 9-11

ABSOLUTE AND RELATIVE MEASURES OF FORECASTING ACCURACY

ACTUAL AND PREDICTED (6) BY (6) INVERSE COEFFICIENTS

	MAD	STPE	RMS	U	UM	US	UC
NAIVE	0.02397	0.07810	0.11977	0.04084	0.02533	0.04831	0.92636
RAS	0.00888	0.02894	0.03365	0.01115	0.00025	0.01366	0.98609
RECRAS	0.03734	0.12164	0.20449	0.07127	0.02996	0.05314	0.91690
PROPVA	0.03145	0.10245	0.16218	0.05576	0.03008	0.04033	0.92959
FRIED	0.00907	0.02956	0.03451	0.01143	0.00020	0.01334	0.98646
RECLAG	0.02464	0.08026	0.10732	0.03517	0.00645	0.04999	0.94356
RASLAG	0.01928	0.06280	0.08341	0.02735	0.00001	0.04919	0.95080
RERALA	0.02801	0.09124	0.12115	0.03960	0.00776	0.05612	0.93612
ALMON	0.13510	0.44010	0.62763	0.21960	0.00334	0.00683	0.98983

RECRAS for UC. The respective lows-highs are 0.00683 to 0.05612 for US and 0.91689 to 0.98984 for UC.

It is evident from table 9-12, that the means of inverse coefficients generated by RAS, FRIED, and RASLAG match that of the actual inverse matrix. The mean is underestimated by RECLAG, RERALA, NAIVE, PROPVA, RECRAS, and ALMON in the amount of 3%, 3%, 6%, 9%, 12%, and 12%. The actual table's standard deviation is overestimated 1%, 1%, 3%, 4%, and 5% by RAS, FRIED, RASLAG, RECLAG, and RERALA. NAIVE, PROPVA, RECRAS, and ALMON underestimate this

TABLE 9-12

MEAN, STANDARD DEVIATION, AND MAXIMUM VALUE

(6) BY (6) PREDICTED INVERSE COEFFICIENTS

ESTIMATION METHOD	MEAN	STANDARD DEVIATION	MAXIMUM VALUE
ACTUAL	0.30697	0.54000	2.0669
NAIVE	0.28791	0.51367	1.92855
RAS	0.30644	0.54393	2.07346
RECRAS	0.27157	0.49286	1.90054
PROPVA	0.27884	0.50743	1.93632
FRIED	0.30649	0.54399	2.07455
RECLAG	0.29835	0.56399	2.17031
RASLAG	0.30669	0.55850	2.16449
RERALA	0.29630	0.56870	2.18618
ALMON	0.27071	0.48814	1.54157

standard deviation by 5%, 6%, 9%, and 10%.

Two of the methods, RAS and FRIED, estimate a maximum value of the coefficients that is equal to the one in the actual table. The maximum values generated by PROPVA, NAIVE, RECRAS, and ALMON are 6%, 7%, 8%, and 25% less than the actual maximum value. RECLAG, RASLAG, and RERALA produce maximum values for inverse coefficients that are 5%, 5%, and

6% above the actual value.

The overall ranking for direct coefficients may be stated as RAS, FRIED, RASLAG, NAIVE, RECLAG, PROPVA, RERALA, RECRAS, and ALMON. The ranking for the case of inverse coefficients is exactly the same except that PROPVA and RERALA exchange places. Here again, all the closeness criterions seem to have reasonable values in both direct and inverse cases, except in some instances of ALMON method.

9.3) REMARKS:

It should be noted that all points and concerns raised in chapter six with regards to interpretation of the results of accuracy measures remain valid here, hence will not be repeated. Having this in mind, and comparing tables 9-1 through 9-12 with one another as well as their counterparts in chapter six, one can infer that these results are in general agreement with those previously obtained. None of the methods provide an estimated table that, in absolute term and uniformly, is very close to the original table. Excluding some subset of coefficients, however, Some of the methods, produce tables that are reasonably close to the survey based table. This conclusion holds true at all levels of aggregation utilized here.

As reported earlier, in some cases the means, the standard deviations, and the maximum values generated by various updating methods match those of the actual tables while in other cases they deviate somewhat from those of the actual tables. However, it seems that the percentage deviations do not follow a pattern and do not have any systematic connection to the level of aggregation. In other words, it appears that the level of aggregation does not affect the central tendency and dispersion of the direct or inverse coefficients in a systematic manner. While the magnitude of the mean and standard deviation are, as expected, increasing with the level of aggregation, the percentage changes are free from any influence exerted by the level of aggregation.

With regards to the accuracy measures, it must be noted that generally the values of MAD statistic increases with aggregation, while STPE, RMS, and U decline as one moves towards more aggregated tables. Thus, it may be inferred that aggregation generally leads to obtaining of a closer estimated table than the case of less aggregation. This implies that if for a researcher, specific industry, commodity, or cell by cell accuracy is not of primal concern and instead, a broad sectoral or overall accuracy is desired, usage of a more aggregated table is advisable. This conclusion, however is based on a limited experiment and may not be generalized without further investigation.

In terms of relative rankings and regardless of the level of aggregation, it appears that RAS and FRIED are consistently the top two performers in both direct and inverse coefficient cases. ALMON, on the other hand, is always ranked at the bottom of the list. The places of other methods are not stable and vary from case to case. The exception is drastic deterioration of RASLAG after aggregation. It should also be noted that the RAS and FRIED methods, while performing very close to one another, are always outperform other techniques by a rather large margin, to the degree that the third place in ranking is invariably a distant third.

CHAPTER TEN

SUMMARY, CONCLUSIONS, FUTURE RESEARCH

"The iron tongue of midnight hath told twelve."

William Shakespeare

"The philosophers have only *interpreted* the world in various ways; The point, however, is to *change* it."

Karl Marx

"And then will begin the rush that will never be checked, the tide that will never turn till it has reached its flood-that will be irresistible, overwhelming-"

Upton Sinclair

This chapter concludes the current endeavor by providing a summary of the project, presenting the major results of the experiments, and suggesting some directions for further exploration and future research.

10.1) SUMMARY:

In the first chapter, the input-output analysis is introduced and its historical roots are briefly traced. Then, the outline and purpose of the current project is mapped out.

Chapter two states the theoretical foundations as well as the underlying assumptions of input-output analysis. The chapter continues by providing the mathematical presentation of input-output analysis along with a typical input-output table. Explanation of this table and the existing interrelationships among its various components concludes chapter two.

Chapter three is devoted to the literature search. The assumption of constancy over time of the I-O coefficients is discussed. The search at this point is primarily focused on the question of intertemporal stability of I-O coefficients. The literature review reveals that most analysts agree on the fact that I-O coefficients do change over time and tables must be constructed via actual surveys to reflect these changes. A consensus among researchers, however, seems to emerge. That is, the solution of actual survey may not be feasible in many cases due to resource constraints imposed on analysts. Furthermore, relatively long time lags exist in construction of survey based tables, which exacerbate the situation to the point that finding of a "shortcut" becomes inevitable. Thus, the focus must be shifted to search for some nonsurvey methods for updating of an actual survey based table. This "shortcut" method must be such that leads to a reasonable estimate of the actual table for a given target year, and achieve this in a faster and less expensive manner. The remainder of this chapter attends to the search for these nonsurvey techniques. Review of the literature on other subjects related to the plan and purpose of the current study is not pursued in this chapter. Instead, they are deferred to their pertinent sections of the project.

The literature search led to selection of several updating methods to be used in the empirical part of the present work. These methods were selected due to their wide acceptability and usage in the field along with their logical plausibility and ease of solution. Explanation of these techniques, providing their mathematical presentation, probing their theoretical and structural foundations, and discussing some related issues is the subject matter of chapter four. The techniques selected in the current undertaking belong to one of the three major categories, biproportional adjustment methods, mathematical programming techniques, and simple naive approaches.

RAS and RECRAS are selected from the family of biproportional adjustment methods. The first one is the familiar RAS procedure and the second one is fundamentally the same, with the exception of its treatment of the value added vector. RAS procedure involves estimation of the square matrix of coefficients and, relying on the fundamental I-O relationships, treats the value added vector as residual. RECRAS method, on the other hand, includes the value added vector in the original matrix and directly estimates the rectangular matrix of technology plus value added vector.

FRIED and ALMON belong to the mathematical programming category. Both of these methods utilize the Lagrangian multiplier optimization technique and the only difference between the two is their proposed minimand. The former employs the minimand suggested by Friedlander, while the latter uses the one suggested by Almon.

PROPVA and NAIVE are members of the simple naive methods of updating I-O tables. The first method assumes that all transactions are proportional to value added. This assumption, in turn, implies existence of a fixed proportion between the usage of any input in production of a given commodity and the amount of labor and capital used in that process. The NAIVE method simply assumes no change in the coefficients during the period under study and applies the base year's coefficients to the target year.

It has been suggested that biproportional procedures may introduce an upward bias into the estimates. The Lagrangian optimization method, on the other hand, does not suffer from such a deficiency. Hence, a two stage process have been proposed. In this procedure, first an estimate is obtained via utilization of the biproportional method, and then, the resultant coefficients are subjected to the Lagrangian optimization technique in order to minimize the distance between the actual and updated matrices. In the present study, RASLAG and RERALA are methods formed from combining RAS and RECRAS with the Lagrangian approach in the manner just described. The minimand used in both cases is the one suggested by Friedlander.

The last technique used, namely RECLAG, is application of rectangularization idea to the Lagrangian approach, using the Friedlander's minimand. In other words, unlike FRIED method the value added vector is not treated as residuals. Instead, this vector has been incorporated into the matrix and is estimated directly.

Thus, utilizing the target year's marginal totals, nine updated matrices of coefficients for the target year may be obtained via application of the aforementioned methods to the actual base year table. This in turn, will yield nine inverse matrices for the target year.

Chapter five describes the data and the implementation process. The actual survey based input-output tables of the Soviet Union for the years 1966 and 1972 are selected. Both tables are expressed in current producers' price and both tables had to be subject to some aggregation in order to be compatible and lend themselves to the desired empirical study. The final versions have (71) by (71) technology matrices. The reasoning behind selection of the Soviet data, choice of current producers' price variant, the particular dates taken, and need for aggregation, is discussed in the first part of chapter five. This part also addresses the question of reliability of the Soviet statistics and its consequences for the present research.

Next section of chapter five is devoted to explanation of the simulation procedure. Using the actual 1966 table as the base year matrix along with marginal totals of the actual 1972 table, and utilization of updating methods selected in chapter four, nine estimates of the 1972 direct and inverse coefficient matrices are obtained. These estimates, then, are compared with their actual 1972 counterparts, thus leading to comparison and ranking of the estimation techniques.

The third section of chapter five concentrates on the methods of matrix comparison and the ways "closeness" is defined in this context. The "partitive" and "holistic" comparisons are described, reasoning is offered as to why a single criterion may not be proper for comparison purposes, and why a "package" of statistical tests is needed.

The last segment of chapter five offers the explanation of the elements of a statistical "package" that is assembled in this project to accomplish the task of matrix comparison and ranking of various updating methods. Utilizing the search of the literature and pointing out the strengths and weaknesses of each closeness test, it is argued why some tests might be useful, hence should be included in the "package," while others do not serve a useful purpose and may be excluded from consideration.

The "package" itself contains Wilcoxon rank-sum test, regression analysis, chi-square test, coefficient of equality, mean absolute deviation, standard total percentage error, root mean square, mean of estimates, standard deviation of estimates, maximum value of estimates, and Theil's U along with its three components UM, US, and UC.

Chapter six begins with explanation of the four stages of matrix comparison, moving from cell by cell contrast of direct matrices to comparing the total outputs. It also provides the reasoning why in this project only the direct and inverse matrices are compared and multiplier as well as total output comparisons are not pursued. Then, the chapter presents the results of the application of selected methods to 1966 table. The balance of chapter six furnishes and discusses the results of comparison of eighteen updated versions of 1972 direct and inverse matrices with the actual 1972 tables. The tests used for this purpose are those included in the statistical "package" mentioned previously.

In the original set of experiments, the non-negativity test indicates that all methods, except ALMON, for both direct and inverse cases produce no or very small (two in this case) number of negative coefficients. The Wilcoxon test at 5% level of significance reveals that, with the exception of ALMON and RECLAG, none of the techniques consistently overestimate or underestimate the coefficients in large numbers. The number of deviations that fall outside of the acceptable range for these methods are zero, one, or two columns. The situation is the same in the case of inverse coefficients, except that the number of columns underestimated by PROPVA increases to seven.

The analysis of coefficient of equality suggests that none of the methods update a large percentage of the coefficients that falls within 5% of their true values. As the acceptance intervals are increased to 10% and 20%, the situation improves somewhat but the numbers still do not seem to be very high. This assessment is further reaffirmed via analyzing the means, standard deviations and the maximum values of the coefficients of equality. Excluding ALMON, which is out of line, the rest of the techniques, including the NAIVE method, perform relatively close to one another. This, more than being a virtue for NAIVE method, is a vice for other techniques. For inverse coefficient the general trend is the same, while the absolute performances improve somewhat. The ranking of the techniques remains the same in both direct and inverse coefficient cases and does not change as acceptance interval widens.

All methods in direct and inverse cases provide statistically significant regression equations as indicated by the F-test at 1% and 0.5% levels of significance. All techniques, except ALMON, show high coefficients of determination and absence of serial correlation as displayed by the Durbin-Watson test at 1% and 5% levels.

Using 1% and 5% level of significance, the t-test indicates that the direct coefficients estimated by all methods, except RECLAG, have an intercept that, in accordance with *a priori* expectation, is statistically not different from zero. The same is true for inverse coefficients, except that NAIVE, PROPVA, and ALMON join the RECLAG in producing an intercept that is different from zero. For the slope coefficient and 1% and 5% levels, estimated direct coefficients match their *a priori* expected value of (1) only in the cases of RAS, FRIED, RASLAG, and RECRAS. The same assertion is valid in the case of inverse coefficients, except that RERALA also provides a slope coefficients of (1).

Testing the null hypothesis of simultaneous equality of intercept coefficient to zero and slope coefficients to one, via utilization of a joint F-test at 1%, leads to rejection of the null hypothesis in all cases except RAS, RASLAG, and RECRAS methods. The same test for inverse matrix results in rejection of the null hypothesis in all but RAS, RASLAG, and RERALA methods. At 0.5% level for direct coefficients, one fails to reject the null hypothesis in cases of RAS, RECRAS, RASLAG, and FRIED. For inverse coefficients RERALA enters the list and FRIED is dropped from the list.

The evidence of regression analysis, then indicates that some of the techniques produce reasonably close estimates of the actual table. The simultaneous satisfaction of the expected values of the intercept and slope coefficients, however, is less frequent. It should be remembered that the distribution of errors may change the results drastically and lead to erroneous conclusion of close estimates when in actuality there is none, or alternatively leads to mistaken conclusion of no similarity while really there is one. This phenomenon is due to the fact that errors of positive and negative magnitude tend to balance each other in the process.

The chi-square test is also performed as a supplemental test of closeness. The test is conducted at 1% and 0.5% level of significance for two sets of classifications of the coefficients, namely one class and forty classes. In the case of one class, for both direct and inverse coefficients, similarity of estimated and actual matrices is concluded. Under forty class case, however, no method is judged to produce a distribution similar to the distribution of actual direct coefficients. For inverse coefficients, on the other hand, all but ALMON demonstrate similarity of distributions.

These somewhat ambiguous results may be explained by the nature of the chi-square test as well as the updating methods. First, comparison of inverse coefficients is a step closer to holistic comparison, hence more likely to demonstrate a similar distribution to the actual table than one can obtain in the case of direct coefficients. Second, existence of very small or zero coefficients can skew the result of chi-square test and lead to misleading inferences. Third, chi-square test fails to capture the structural differences in matrices. It is incapable of addressing the "relocation" problem associated with the updating procedure, hence may lead to erroneous conclusions.

MAD, STPE, Mean of the estimates, Standard deviation of the estimates, Maximum value of the estimates, and Theil's U along with its components UM, US, and UC are also used as measures of closeness of updated matrices to the actual one. The results are again mixed. MAD values seem to be low, even though there is no systematic way to substantiate this assertion. In absolute term, the direct coefficients have smaller MAD values than the inverse coefficients. Excluding ALMON, the STPE values may be considered marginally acceptable for direct coefficient and, with sizeable improvement, reasonable for the inverse coefficients. The same may be inferred for U values as well. The RMS values in absolute term are smaller for the direct coefficient cases. With the exception of ALMON, these values may be judged as reasonable.

With regards to the mean and standard deviation of the coefficients in both direct and inverse cases all methods except ALMON produce reasonably close statistics to those of the actual table. Almost similar pattern is observed for the maximum value of the coefficients. In direct case RECRAS, NAIVE, and ALMON give values that substantially deviates from that of the actual table. For inverse coefficients this is true only in ALMON's case. Values of UM, US, and UC are all very close to those of the actual table. This in itself, of course, is not striking, since it may simply be a good distribution of bad results. The overall ranking of the methods may be stated as RAS, RASLAG, FRIED, RECRAS, RERALA, NAIVE, PROPVA, RECLAG, and ALMON.

Chapter seven takes on the question of incorporation of exogenous information into the estimation process. A brief literature search on the subject is performed and it is concluded that most researchers believe that such incorporation will improve the results, hence this inclusion is advisable and sometimes even necessary. To investigate this claim, RAS and FRIED methods are selected to be "modified" via incorporation of exogenous data and analysis of the resultant tables. The NAIVE or constant coefficient case is also included in this list for comparative purposes. An additional modified method, namely RESMIN, is also included for its simplicity and plausibility.

In the next step, all these methods are explained and their pertinent mathematical relationships are offered. The focus of the chapter is turned to selection of criterions for determining which coefficients should be estimated exogenously and which ones should be left for estimation through the updating methods. To this end, three criterions are suggested in the chapter, i.e., "KEY", "BIG", and "MIP".

The first criterion refers to a set of coefficients that, for variety of reasons, may assume a key position in a given economy, thus warranting their exogenous estimation. The second criterion, following the common wisdom of the literature, is selection of a certain percentage of the largest coefficients in the table for separate estimation. The last criterion is the result of an algorithm that is constructed via utilization of the concepts of "norm" and "condition number" of a matrix and selects the most important coefficients in a matrix for exogenous estimation, i.e., those coefficients whose changes will have the largest impact throughout the table. Explanation and logic of these criterions is the subject matter of the rest of the chapter.

Chapter eight explains the logic and process of modification of tables through incorporating exogenous data. The modified tables are procured via application of the three selection criterions mentioned above to RAS, FRIED, and NAIVE methods. Thusly, three modified variants of each selected updating technique are obtained. These are MDRASKEY, MDRASBIG, MDRASMIP, MDLAGKEY, MDLAGBIG, MDLAGMIP, MDNAVKEY, MDNAVBIG, and MDNAVMIP.

The "key" sectors selected for this purpose are the energy sectors in the Soviet Union, which amounts to 497 coefficients. Then, for the sake of compatibility, 497 largest and 497 most important coefficients are chosen via employment of "BIG" and "MIP" criterions. In the modification process, the actual values of the 497 coefficients selected by various criterions are taken from the 1972 table and treated as if they were exogenously estimated. The balance of each table is estimated via employment of an updating technique, and then incorporated with the "exogenously determined" coefficients to arrive at the final estimated table for 1972. Thus, nine estimates of the target year's direct and inverse matrices are at hand. Adding the RESMIN method to this list, total of ten modified direct coefficient matrices and ten modified inverse coefficient matrices are acquired. These estimated matrices and their pertinent inverses, then, are subjected to the "package" of closeness tests, which were introduced in chapter five, and the results are reported and discussed in the remainder of chapter eight.

The number of negative coefficients generated in modified cases of FRIED are increased although these numbers are still negligible. The Wilcoxon test at 5% shows that for both direct and inverse cases no method overestimates or underestimates more than three columns. Analyzing the coefficients of equality, it can be observed that the number of coefficients estimated within 5%, 10%, and 20% of the actual values are increased. Discounting the 497 exogenously determined coefficients, however, reveals that these methods actually estimated less number of coefficients within the desired intervals than their unmodified counterparts. This may be viewed as an evidence in support of proposition that inclusion of exogenously determined cells actually worsens the estimates of the remaining coefficients.

As the acceptance intervals widens, the number of estimations within the intervals increases, but the total number still remains relatively low. The ranking also varies with different acceptance intervals. The estimates have wide range of variation as is evident from the large deviation of their means, standard errors, and maximum values. In the inverse cases, the situation is somewhat different. Taking into account the 497 exogenously determined coefficients, it can be seen that "BIG" and "MIP" variants of RAS and FRIED outperform their unmodified analogues. The situation remains the same as acceptance intervals widens, and at 20% interval MDNAVBIG and MDNAVMIP also join the rank of the methods who outperform their unmodified twins. Regardless of the error intervals, the ranking of the techniques remains rather stable. The variation in the estimates, although less than the case of direct coefficient, is still large as is evident from the means, standard deviations, and the maximum values of the coefficients of equality.

All methods exhibit a statistically significant regression lines, as determined by F-test at 1% and 0.5% levels, with high coefficient of determination. At 1% and 5% level, Durbin-Watson test points to absence of serial correlation for direct coefficients. For inverse coefficients, however, MDNAVMIP and MDNAVBIG show serial correlation at both 1% and 5%. MDNAVKEY and MDRASMIP demonstrate the serial correlation only at 5% level. At 1% level and for both direct and inverse coefficients, all methods except variants of NAIVE and MDNAVKEY, have zero intercepts. At 5%, only the three NAIVE cases do not have intercepts equal to zero.

For slope coefficient, and in direct cases, at 1% no method except MDLAGMIP, MDRASMIP, and MDRASKEY produces slope of one. At 5%, MDRASKEY is eliminated from this list. For inverse cases and at 1% MDLAGKEY, MDLAGMIP, and MDRASKEY have a slope that is statistically not different from one. At 5% MDRASKEY is excluded from the list. Comparison of these results with those of unmodified cases, reveals that incorporation of exogenous data leads to deterioration of regression lines.

Joint F-test for direct coefficients and at 0.5% shows that only MDLAGMIP and MDLAGKEY meet the *a priori* expectation. At 1%, MDLAGKEY is removed from the list. For inverse coefficients and at 1%, MDLAGKEY, MDLAGMIP, and MDRASBIG pass the test, and as the significance level is changed to 0.5%, MDRASKEY will be added to this roll. Thus, again the results suggest that inclusion of exogenous data in most cases causes deterioration of regression line.

The chi-square test for forty classes direct coefficients, and at 0.5% level of significance, shows absence of similarity of distribution between estimated and actual matrices in all but the cases of MDLAGKEY, MDLAGMIP, and MDLAGBIG. Changing the level of significance to 1% will discard the MDLAGKEY from the list. In the case of inverse coefficients, however, at both levels of significance all tests indicate existence of similarity among estimated and actual matrices. These results demonstrate that the estimated inverse coefficients are closer to their benchmark counterparts than are the estimated direct coefficients to theirs. It is also evident that there is some improvement in the results for estimated direct coefficients in modified versions. In the case of one class chi-square test, at both levels of significance, and for both direct and inverse coefficients, one fails to reject the null hypothesis of similarity of distributions among estimated and actual matrices.

The means, standard deviations, and maximum values of the coefficients in both direct and inverse cases demonstrate a similar central tendency and dispersion to the actual data. It is also evident that these statistics experience improvement over their unmodified counterparts.

MAD, STPE, RMS, and U values for direct and inverse coefficients appear to be acceptably low. It is clear that the inverse cases are closer to their actual values than are the direct coefficients. UM, US, and UC in both cases show very good results.

Compared to the unmodified cases, the "KEY" variant of RAS, NAIVE, and FRIED in both direct and inverse instances demonstrate some improvement over the unmodified ones. The "MIP" and "BIG" variants of the same techniques show substantial improvements over the unmodified cases. The improvement is particularly impressive for the "BIG" variants. Results also point out that modifications, specially through the "BIG" criterion, turn the NAIVE method into a very viable alternative for updating purposes. The overall ranking of the updating techniques in the case of direct coefficients may be stated as MDRASBIG, MDLAGBIG, MDNAVBIG, MDRASMIP, MDLAGMIP, MDNAVMIP, MDRASKEY, MDLAGKEY, MDNAVKEY and RESMIN. For inverse cases the ranking is the same, except RESMIN and MDNAVKEY exchange places.

In many instances, constraints imposed on researchers mandates usage of less aggregated tables. The aggregation and its consequences have long occupied the minds of I-O researchers. Discussion of these concerns, associated problems, and proposed solutions are beyond the scope of the present project and are deferred to a later date. A brief encounter with the issue, however, may be in order. Chapter nine is an attempt to conduct a preliminary investigation of the effects of aggregation on the performance of the updating techniques and test of stability of coefficients.

To accomplish the task, three levels of sectoral details of the original matrices are produced. First level of aggregation reduces the matrix to a (35) by (35) one. The second and third level transform the matrix into (16) by (16) and (6) by (6) matrices respectively. In the next phase, these aggregated matrices are updated via utilization of the original nine updating techniques. Accordingly three sets of estimated direct and inverse coefficients are acquired. Each set, of course, contains nine estimates of direct and nine estimates of inverse coefficients.

In the last part of chapter nine, the estimated direct and inverse matrices of all three levels of aggregation are subjected to selected closeness tests. The results are compared with one another as well as those of the "original" (71) by (71) matrix, and the effects of aggregation is discussed.

Investigation of the results points discloses that all methods, except ALMON, at all three levels of aggregation and for both direct and inverse coefficients cases produce means and standard deviations that are fairly close to those of the actual tables. Excluding the ALMON case, the maximum values estimated via different methods show estimates that range from very close to reasonably acceptable. The means and the standard deviations of the estimated direct coefficients become larger, and their maximum values become smaller with aggregation. Their percentage deviations from their actual counterparts, however, remain fairly stable. The situation for inverse coefficients are basically the same, except that the percentage deviations are slightly increased for some methods, and the maximum values of some techniques do not always decreases with aggregation.

Due to its poor and unsystematic behavior, ALMON method is excluded from further discussion. For the remaining techniques, it may be said that in both direct and inverse coefficient cases the values of MAD statistics are an increasing function of the level of aggregation. These values, in absolute term, seem to be reasonable within the context of each aggregated table. The values of STPE, RMS, and U statistics, with some sporadic exception, display improvements with aggregation. The exceptions for the most part are in the cases of RASLAG, and RERALA methods. Values of the components of U are all appear to be good and show no systematic correlation to the level of sectoral detail.

Thus, it may be inferred that aggregation generally improves the estimates, and in some cases they are very close to the actual tables. The estimated inverse coefficients, in general, are closer to the benchmark tables than the estimated direct coefficients. Overall for both direct and inverse coefficients, and at all levels of aggregation, it can be said that ALMON method is always ranked at the bottom of the list. RAS and FRIED, on the other hand, always outperform other techniques with a comfortable margin and occupy top two positions. With few exceptions, RECRAS and NAIVE come next in line. The places of other techniques is not stable and varies with level of aggregation. With the exception of a rather drastic deterioration of RASLAG method's performance, the results obtained from aggregated tables meet a *priori* expectations, and are in general agreements with those obtained in the original less aggregated experiments.

10.2) CONCLUSIONS:

The experiments conducted in chapters six, eight, and nine provided results that were reported in those chapters and are summarized in the previous section. In this section, the major findings of the current research are reiterated and some conclusions are drawn.

Prior to pronouncement of any conclusion, it must be noted that the experiments conducted in the present endeavor are limited in nature and scope. Thus, the results may not be universally applicable, spatially or intertemporally. For it is always possible to obtain good results at one period or area and less desirable results at another. Furthermore, the time span covered by the experiments are rather short, hence may not be valid for the longer periods. Finally, the Soviet economy during the period under study did not experience drastic structural or technological metamorphosis. This fact, in turn, tends to stabilizes the input-output coefficients. With these points in mind, then, the findings may be presented.

It is evident that the I-O coefficients, even for a relatively stable economy and short time span covered, do change over time. Thus, reliance on the most recent survey based table for forecasting purposes may be misleading. This assertion is fortified by the fact that several updating techniques generate estimates superior to that of the NAIVE method. This may also be a testimony to the usefulness and necessity of such updating procedures.

At the first stage of experiments, none of the estimation methods seem to be able to replicate the benchmark table in partitive sense. Existence of large number of very small coefficients in the original table may be a major contributor to this phenomenon. Some sub blocks of coefficients, however, appear to be accurately estimated by some of the techniques.

An appreciable improvement is obtained when one moves to comparison of inverse coefficients, which indicates better estimates are obtained in operational (holistic) sense. This in turn implies that the updating techniques are more useful if holistic, rather than partitive, accuracy is desired. RAS and FRIED methods are the best performers among all methods utilized here. FRIED method's inability to guaranty generation of non negative coefficients puts this technique at some disadvantage. Of course, non negativity conditions may be imposed on the model at the formulation stage. This imposition, however, elevates the method into quadratic programming domain, with much heavier computational demand. Although the NAIVE method does not generate very accurate results, it is not very far behind the top performing techniques. This is true specially for some subset of coefficients, and particularly as upper and lower boundaries of error tolerance intervals are widened. The implication of this last observation could be relative stability of a subset of input-output coefficients.

Further exploration of the results reveal that, incorporation of exogenous data as prior information leads to improvement of the estimates. The improvements are particularly substantial when "BIG" criterion is utilized for selection of exogenously determined coefficients. MDRASBIG and MDLAGBIG are particularly impressive. Also interesting is good estimate obtained via MDNABIG, which again is an indication of relative stability of some sub block of I-O coefficients. Therefore, it may be asserted that the large coefficients exert the most influence on the table, hence increase in their accuracy improves the overall accuracy of the estimated tables. Reasonable estimates are also obtained through modified versions when "MIP" criterion is employed for selection of exogenously determined coefficients. This, again, can lead one to conclude that relatively small number of coefficients exert heavy influence on the entire table.

There are indications that incorporation of prior information in the updating process, while improving the overall accuracy of the table, causes deterioration of the estimates of the remaining coefficients. This phenomenon is probably due to the updating process and very small initial value of some of the coefficients.

Analysis of the results for three levels of sectoral details and comparison of these results with those obtained in the original set of experiments, indicates that some very good estimates are obtained in aggregated cases. It is evident that as level of aggregation increases, the estimates become closer to their benchmark counterparts. This is particularly true in the cases of RAS and FRIED methods. Performance of NAIVE method also seems to be a direct function of level of aggregation. These results may be interpreted as an indication of higher relative stability of coefficients as the tables are aggregated.

Thus, it appears that in the absence of actual survey based input-output tables, updating methods can be of tremendous help to researchers, policy makers, and all other users of I-O tables. These tables are particularly helpful if holistic, instead of partitive, accuracy is sought. Inclusion of exogenous data, if available, and movement to a more aggregated table, if possible, substantially improve the results and enhance the value of the table to the users.

Nothing, however, will replace the actual painstakingly constructed survey based table. Consequently, If feasible, it is recommended that I-O tables be adopted as an integral part of national income accounting practice and be constructed on an ongoing basis. In the construction phase, however, special attention must be paid to particular coefficients that exert the most influence on the entire table. These coefficients may be selected via "BIG," "MIP," or some other suitable criterion. In the interim, updating techniques such as RAS, FRIED, or even NAIVE, specially if they are modified via incorporation of appropriate exogenous data, may be utilized. The recommendation is particularly relevant since the modified updating methods are capable of producing rather accurate estimates for the short and medium runs. This recommendation is further validated in light of the fact that in most cases, holistic accuracy and impact analysis are of primary concern to the I-O analysts.

10.3) **FUTURE RESEARCH**:

The vast topic of input-output analysis was barely touched in the current endeavor. There are many avenues that may be pursued by taking a que from one or more of the points alluded to in the previous pages. Few possibilities, however, are directly connected to and logically follow the objectives and aspirations of this project, and readily lend themselves to further research. Thus, to optimize the initial resource investment, these possibilities should constitute the immediate continuation of the current research.

The survey based 1956 input-output table of the Soviet Union is already available and the survey based 1977 table is in process of being published in the West. Selection of various combinations of the four survey based tables, covering twenty two years, will be the first step in continuation of the line of research developed in this project. This pursuit should shed some light on the long run performance of the updating methods, and at the same time can help to either substantiate and solidify, or repudiate, the results obtained here.

The application of the same methods to another set of available I-O table, such as that of the United States, The Netherlands, Japan, etc., can further corroborate or reject the results obtained in this project.

In this study, aggregation and modification were conducted separately. A logical continuation of the same line of thought would be combination of these two steps together and analysis of the results in search of gain in accuracy of the updates.

Detail analysis of most important parameters ("MIP") and the largest coefficients ("BIG") is also needed. It is interesting to see how many of the coefficients are shared by these two selection criterions, and to investigate the possibility of a systematic connection between the two sets. This search can illuminate the path for search of the most efficient criterion for selective targeting of coefficients for exogenous estimation.

Next step might be in the direction of improving the statistical "package" used for test of closeness of two matrices. This effort may lead to elimination of redundant tests and addition of supplementary ones, with the end result of improvement in performance of the "package."

Additional updating methods may be substituted for poor performers in this project, to explore the possibility of finding a superior updating technique.

Aggregation question was briefly encountered here. A more thorough and detailed experiments with aggregation could provide many useful insight into the behavior of the I-O coefficients. There were some indication that aggregation increase the stability of the coefficients. Future research may be directed towards the analysis of the effects of various levels of aggregation on this stability, and probing the behavior of different sectors at various levels of sectoral detail.

The data used in this study is expressed in current producers' prices. This choice was made to follow the results of prior experiments reported by researches in the field, that suggest higher degree of coefficient stability for tables expressed in current producers' price. In the subsequent studies, the experiments conducted here may be applied to other variants of the same data to test the updating methods' performance as well as to test the assertion that coefficients of tables expressed in current producers' price are more stable than other price versions.

The aforementioned lines of pursuit are only few from a long list of possibilities. These suggestions do not even enter the dominions such as structural analysis of the Soviet economy and its evolution, theoretical investigation of the foundations and assumptions of input-output analysis, theoretical probe of the updating methods and their economic interpretation, etc. Coverage of all or even most of the suggestions and possibilities, however, is far beyond one person's ability. Most of these points are stated more as a suggestion than as an agenda. They are proposed mainly for future researchers, rather than the present one. For the quest is infinite and one person's ability and allotted time agonizingly finite.

APPENDIX A

WILCOXON NON-PARAMETRIC TESTS

FOR

ORIGINAL AND MODIFIED ESTIMATES

DIRECT AND INVERSE COEFFICIENTS

FIGURE A-1

WILCOXON NON-PARAMETRIC TEST FOR DIRECT NAIVE

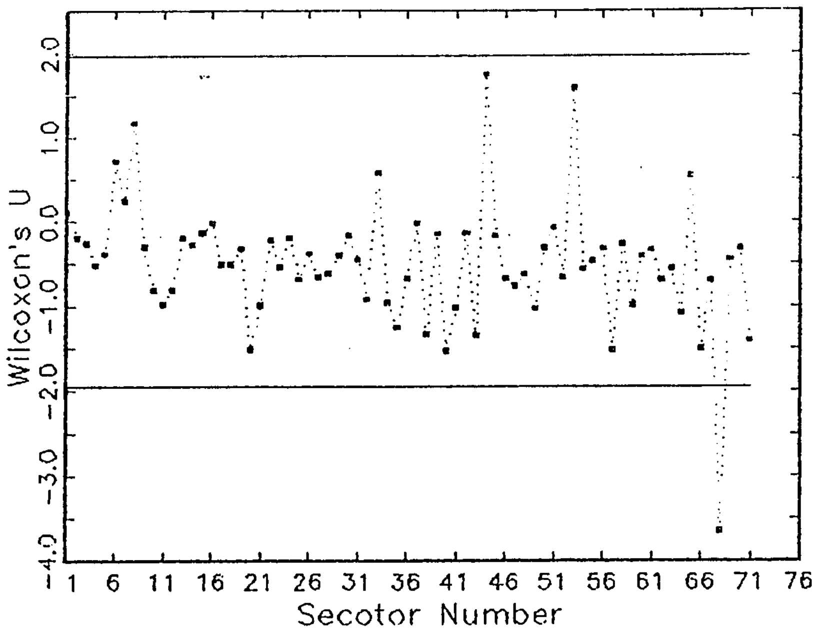


FIGURE A-2

WILCOXON NON-PARAMETRIC TEST FOR INVERSE NAIVE

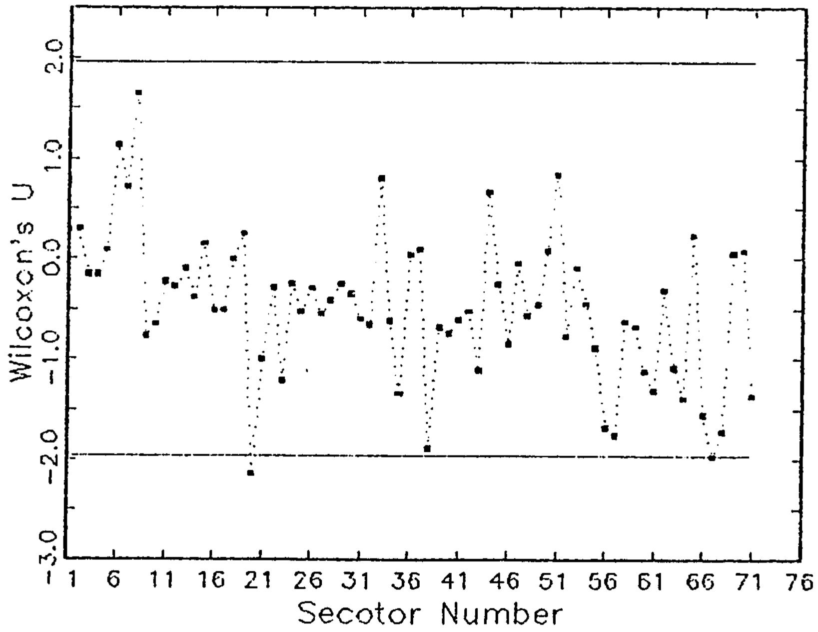


FIGURE A-3

WILCOXON NON-PARAMETRIC TEST FOR DIRECT RAS

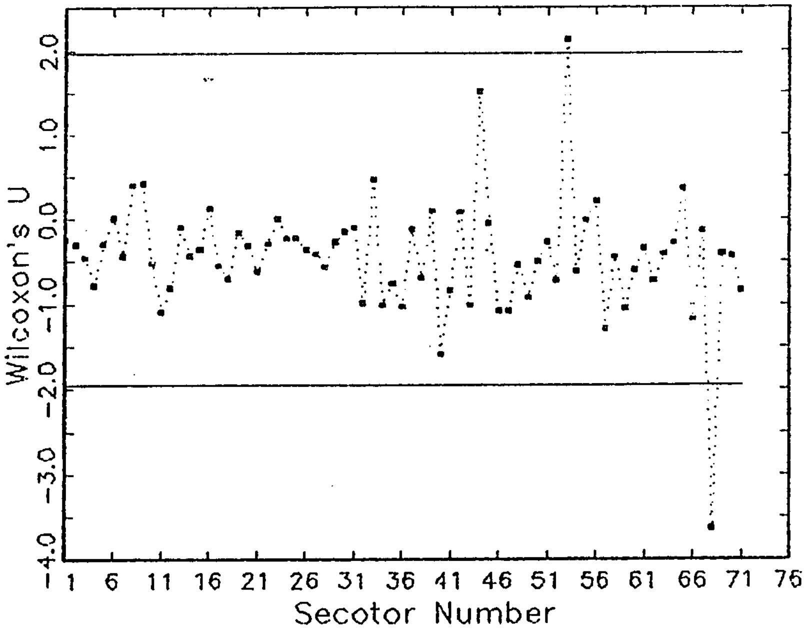


FIGURE A-4

WILCOXON NON-PARAMETRIC TEST FOR INVERSE RAS

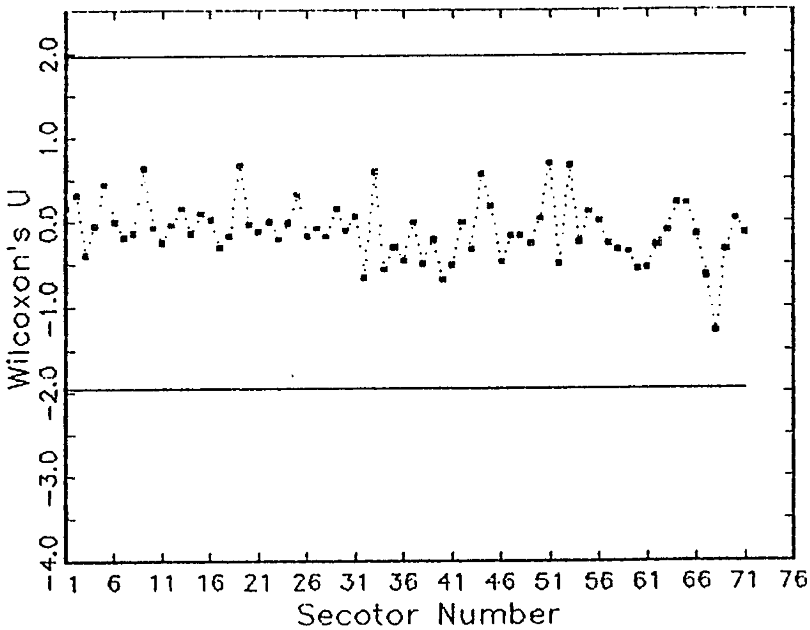


FIGURE A-5

WILCOXON NON-PARAMETRIC TEST FOR DIRECT RECRAS

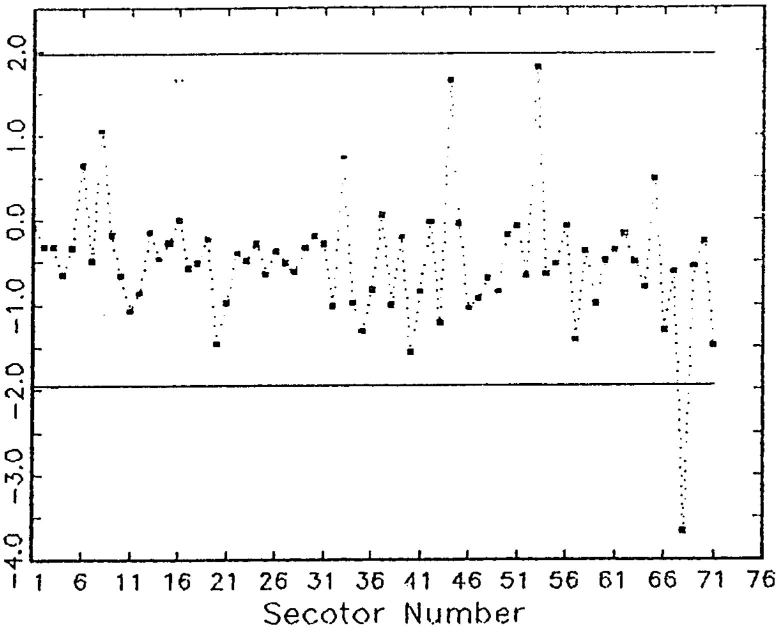


FIGURE A-6

WILCOXON NON-PARAMETRIC TEST FOR INVERSE RECRAS

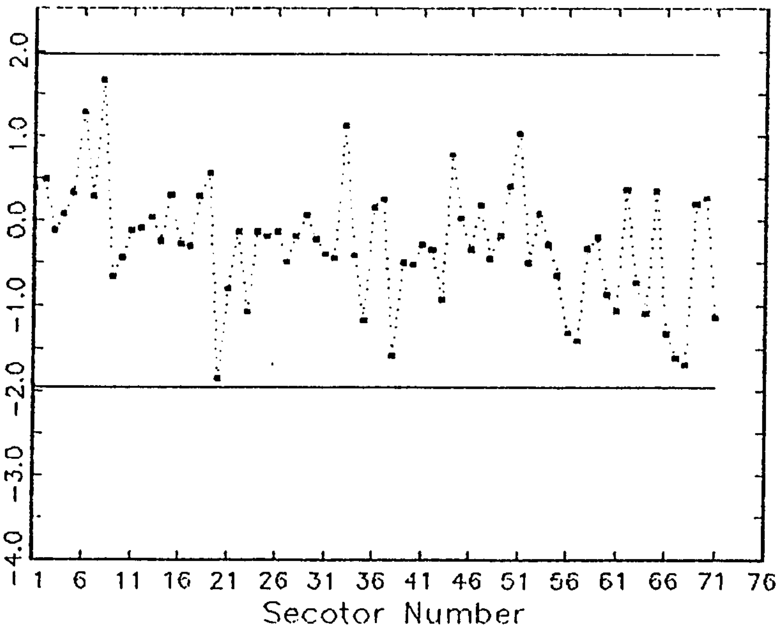


FIGURE A-7

WILCOXON NON-PARAMETRIC TEST FOR DIRECT PROPVA

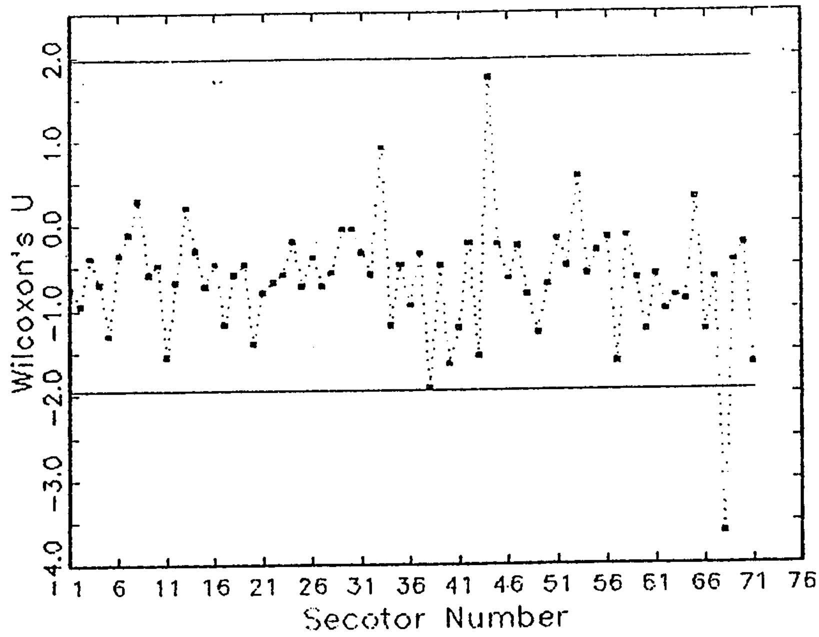


FIGURE A-8

WILCOXON NON-PARAMETRIC TEST FOR INVERSE PROPVA

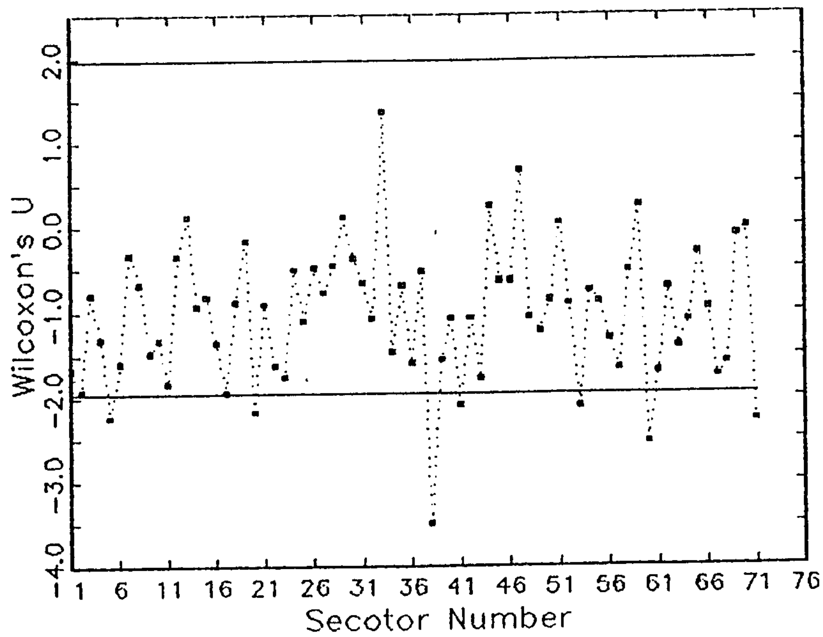


FIGURE A-9

WILCOXON NON-PARAMETRIC TEST FOR DIRECT FRIED

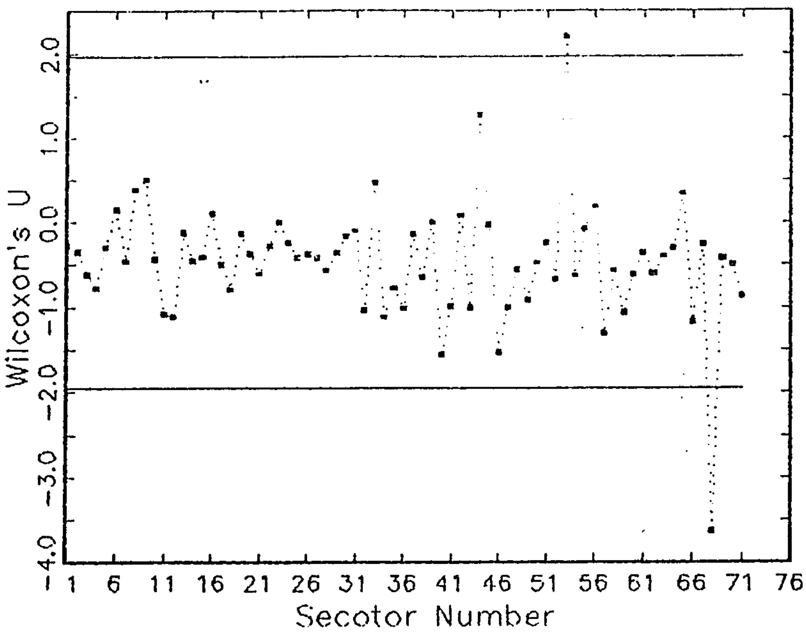


FIGURE A-10

WILCOXON NON-PARAMETRIC TEST FOR INVERSE FRIED

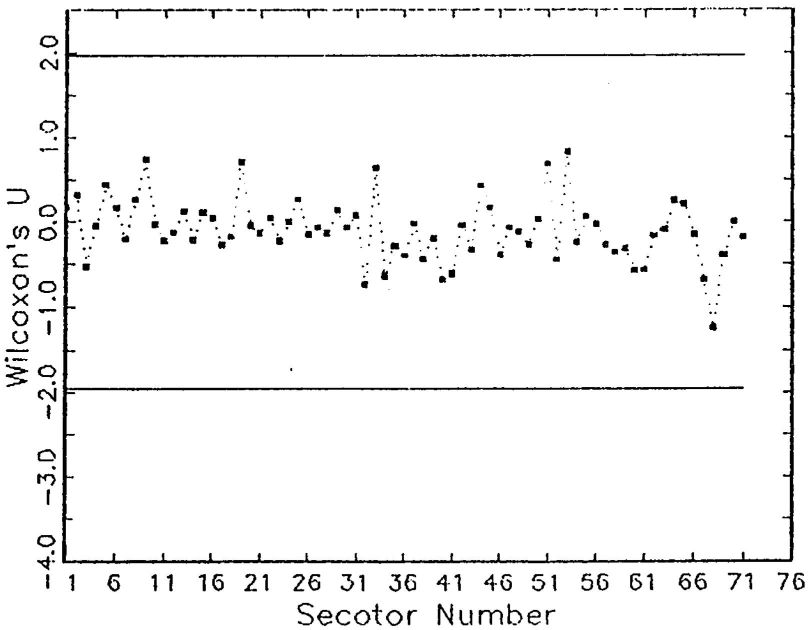


FIGURE A-11

WILCOXON NON-PARAMETRIC TEST FOR DIRECT RECLAG

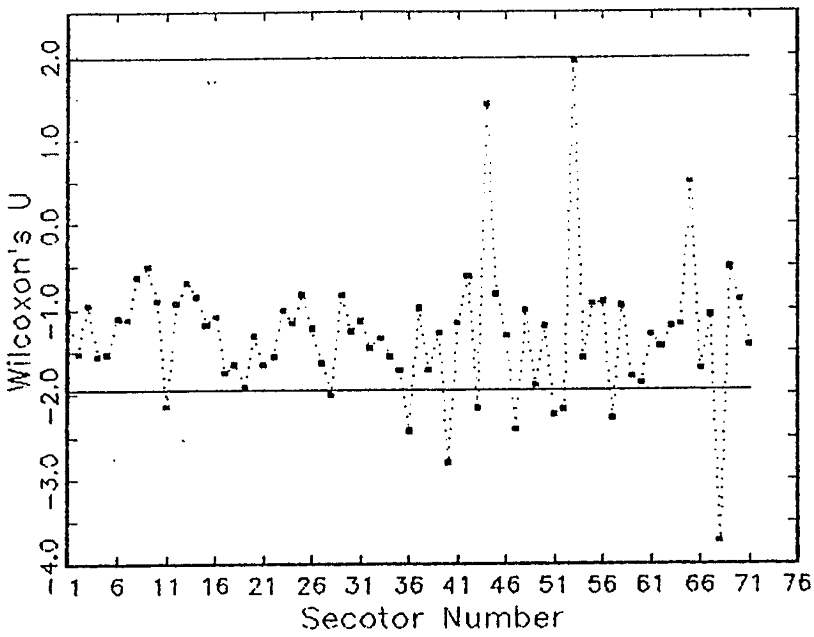


FIGURE A-12

WILCOXON NON-PARAMETRIC TEST FOR INVERSE RECLAG

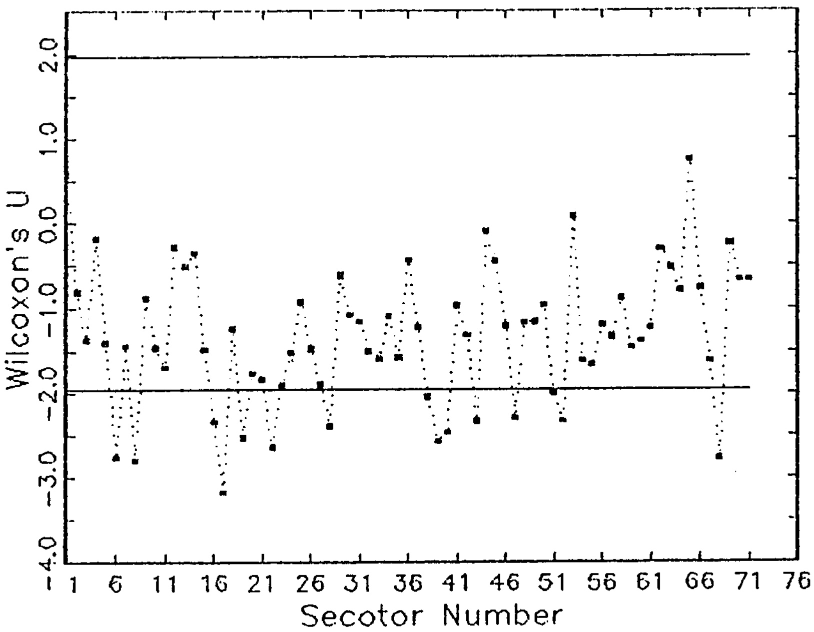


FIGURE A-13

WILCOXON NON-PARAMETRIC TEST FOR DIRECT RASLAG

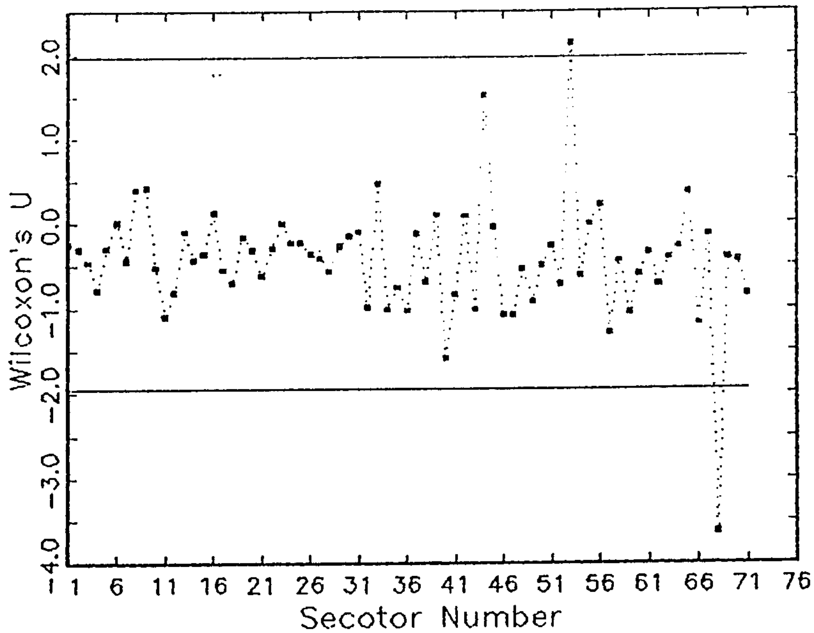


FIGURE A-14

WILCOXON NON-PARAMETRIC TEST FOR INVERSE RASLAG

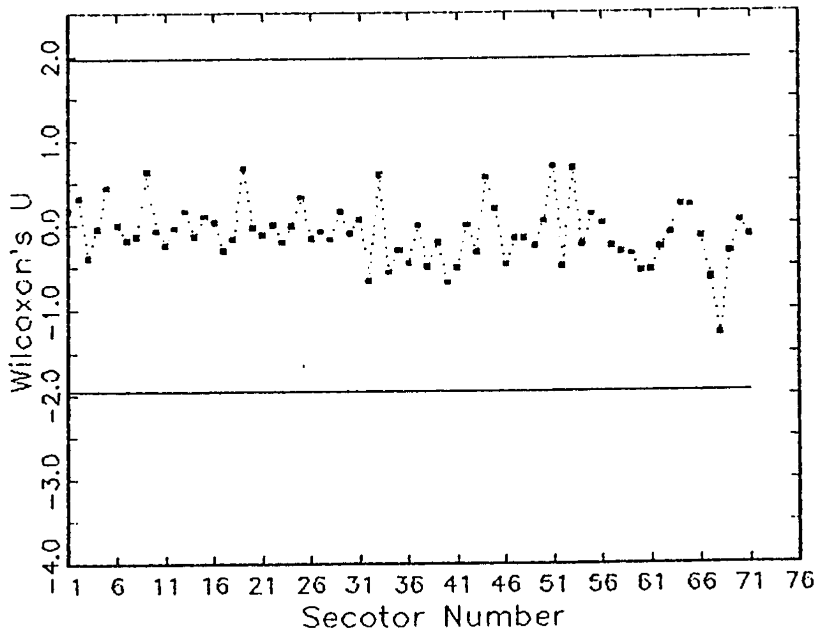


FIGURE A-15

WILCOXON NON-PARAMETRIC TEST FOR DIRECT RERALA

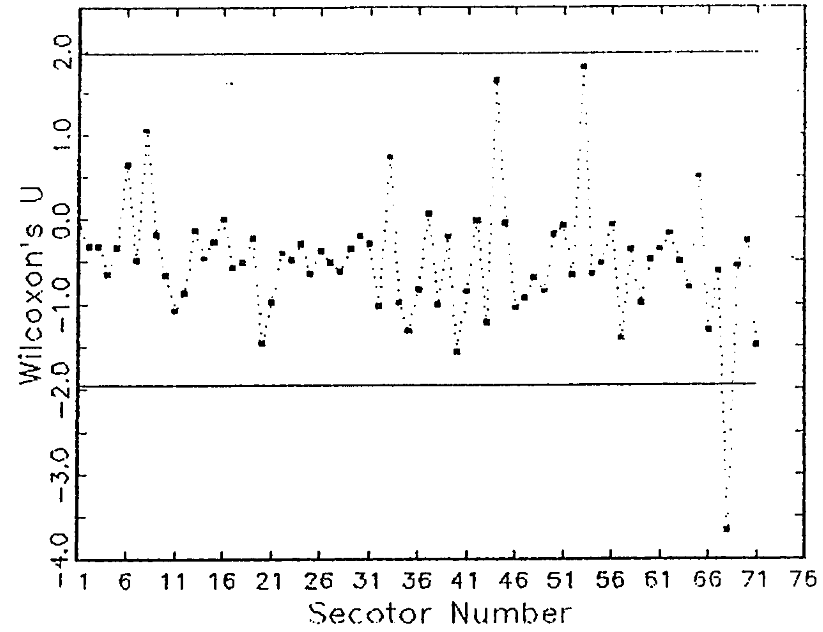


FIGURE A-16

WILCOXON NON-PARAMETRIC TEST FOR INVERSE RERALA

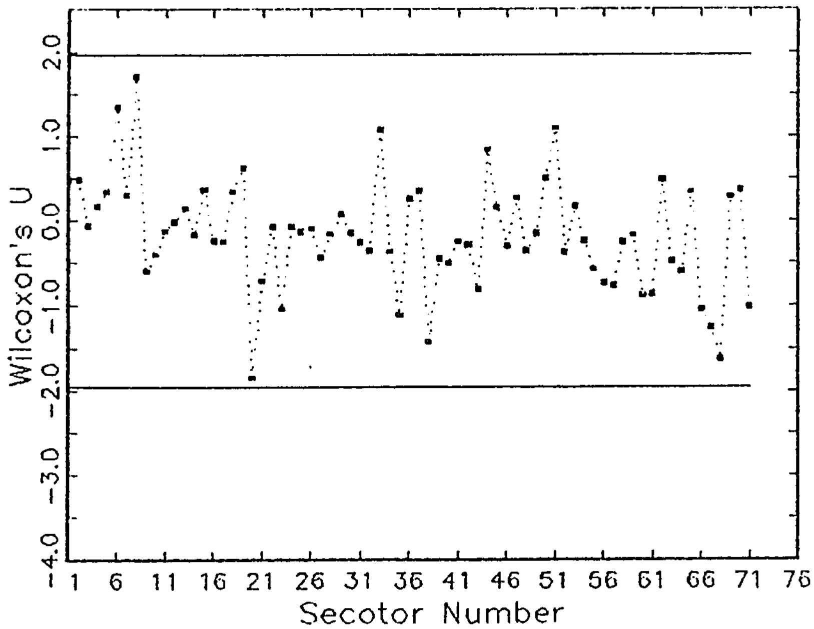


FIGURE A-17

WILCOXON NON-PARAMETRIC TEST FOR DIRECT ALMON

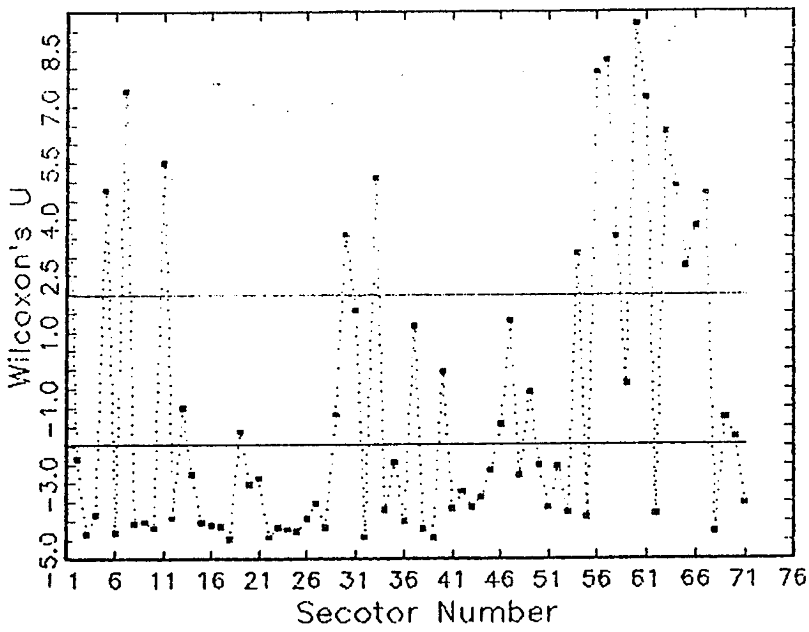


FIGURE A-18

WILCOXON NON-PARAMETRIC TEST FOR INVERSE ALMON

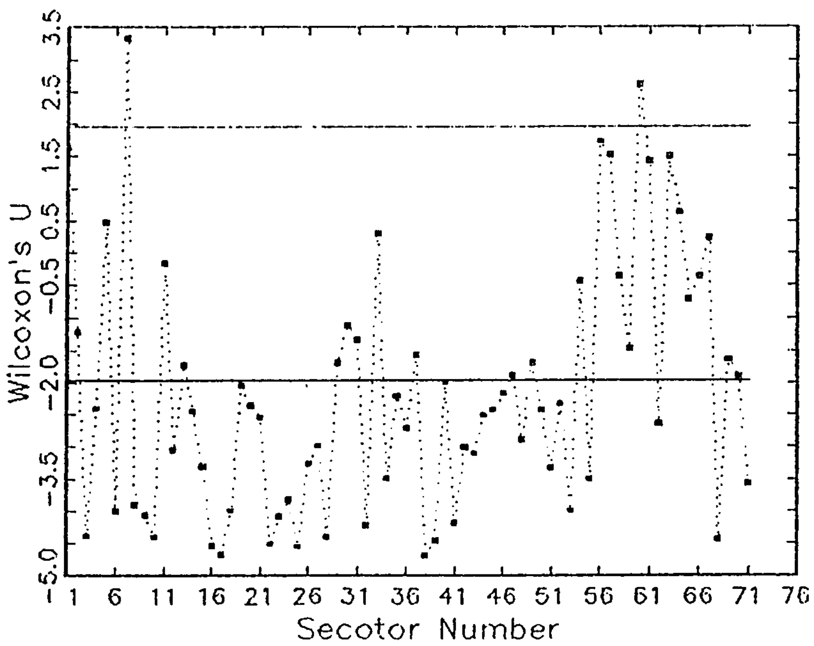


FIGURE A-19

WILCOXON NON-PARAMETRIC TEST FOR DIRECT MDNAVKEY

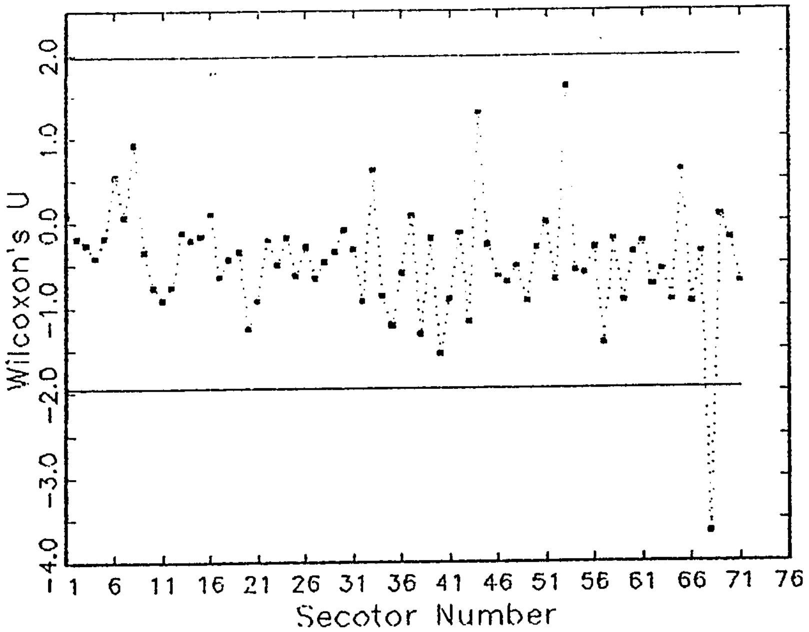


FIGURE A-20

WILCOXON NON-PARAMETRIC TEST FOR INVERSE MDNAVKEY

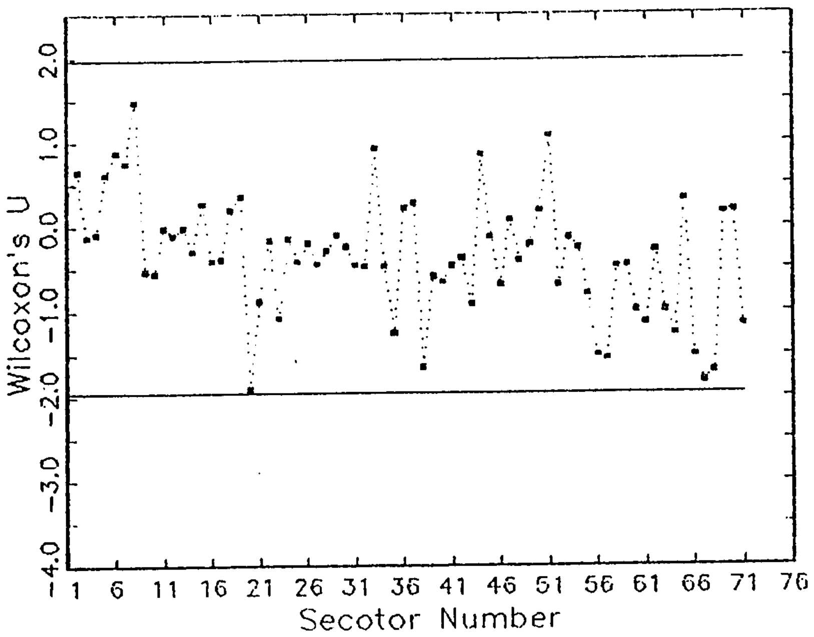


FIGURE A-21

WILCOXON NON-PARAMETRIC TEST FOR DIRECT MDNAVBIG

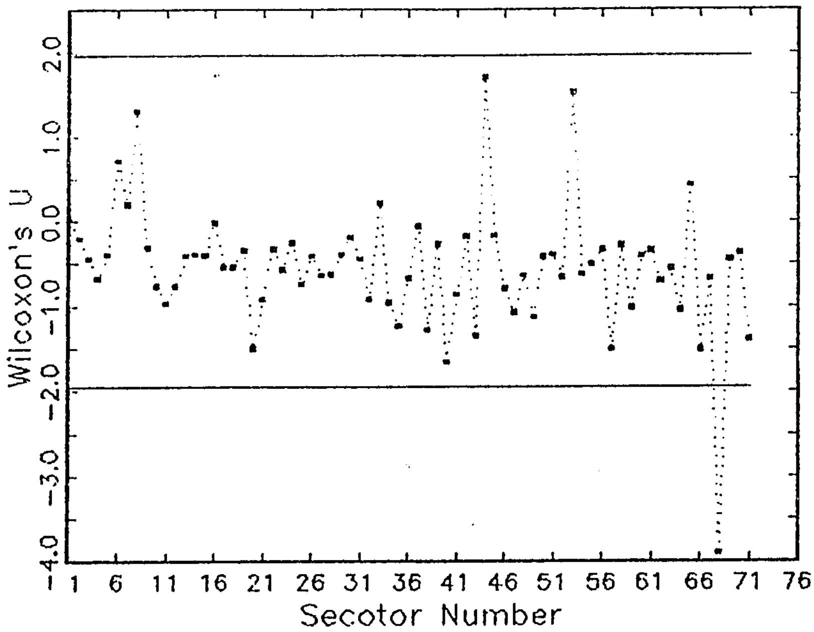


FIGURE A-22

WILCOXON NON-PARAMETRIC TEST FOR INVERSE MDNAVBIG

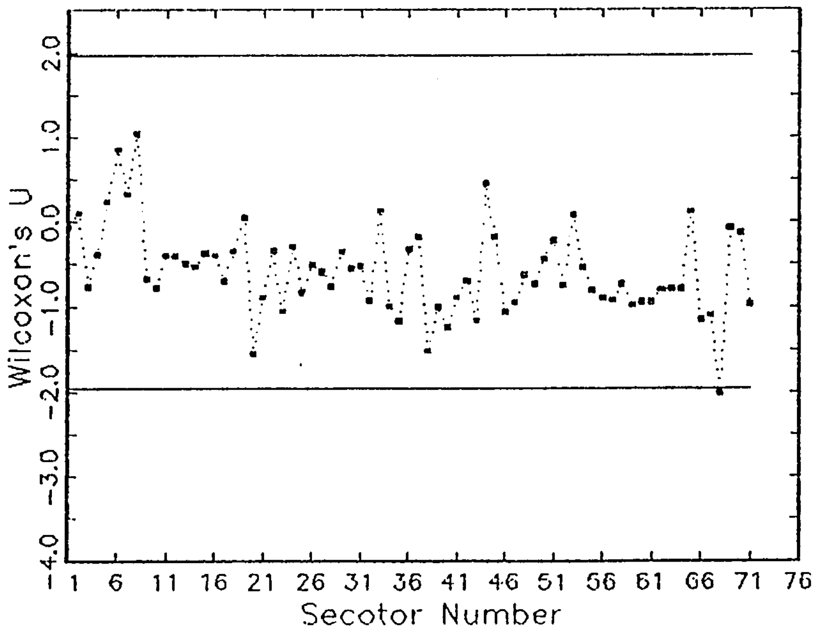


FIGURE A-23

WILCOXON NON-PARAMETRIC TEST FOR DIRECT MDNAVMIP

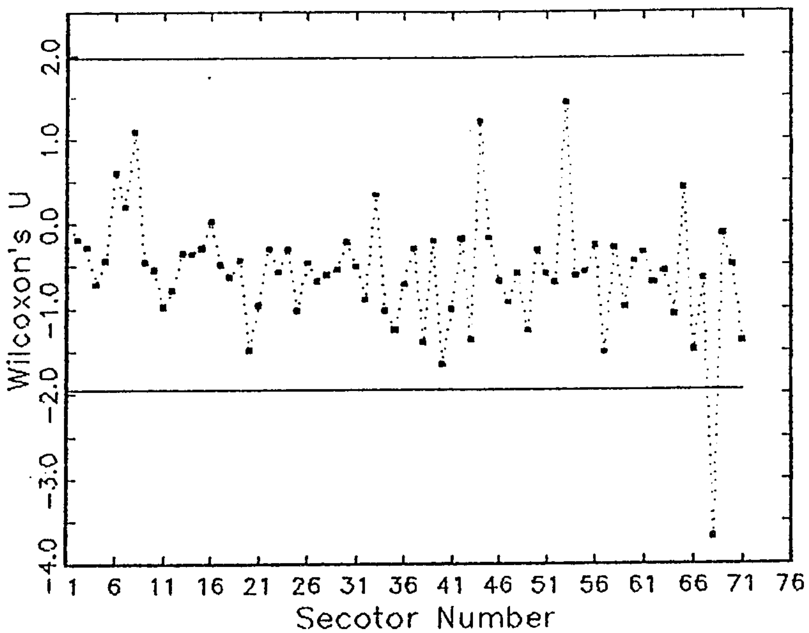


FIGURE A-24

WILCOXON NON-PARAMETRIC TEST FOR INVERSE MDNAVMIP

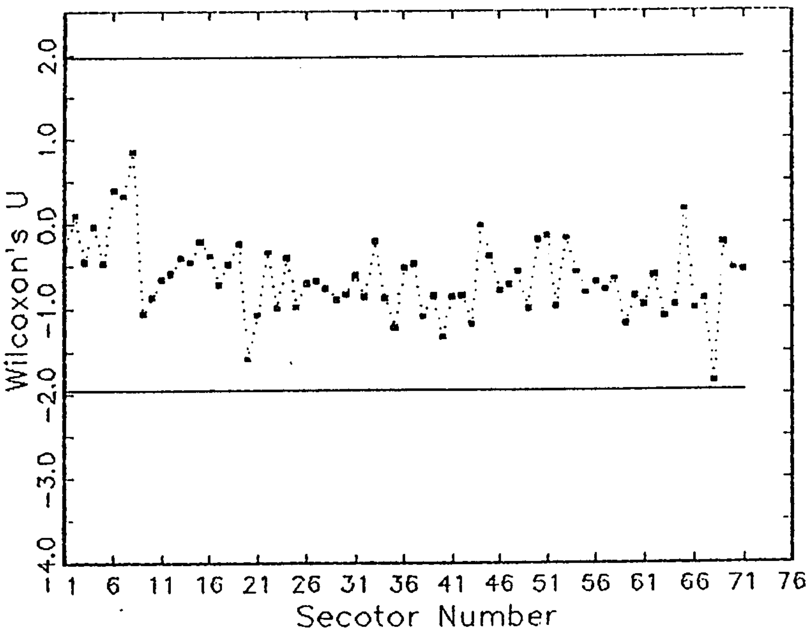


FIGURE A-25

WILCOXON NON-PARAMETRIC TEST FOR DIRECT MDLAGKEY

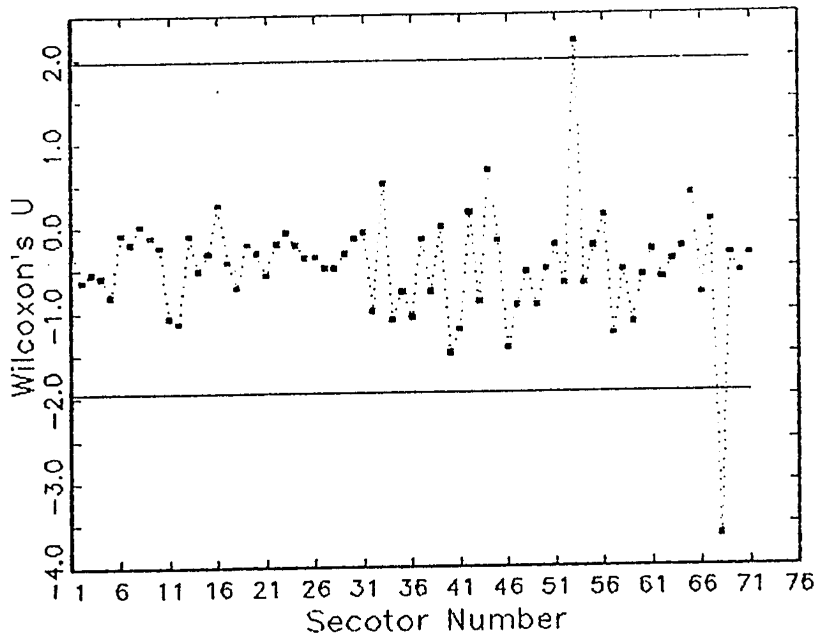


FIGURE A-26

WILCOXON NON-PARAMETRIC TEST FOR INVERSE MDLAGKEY

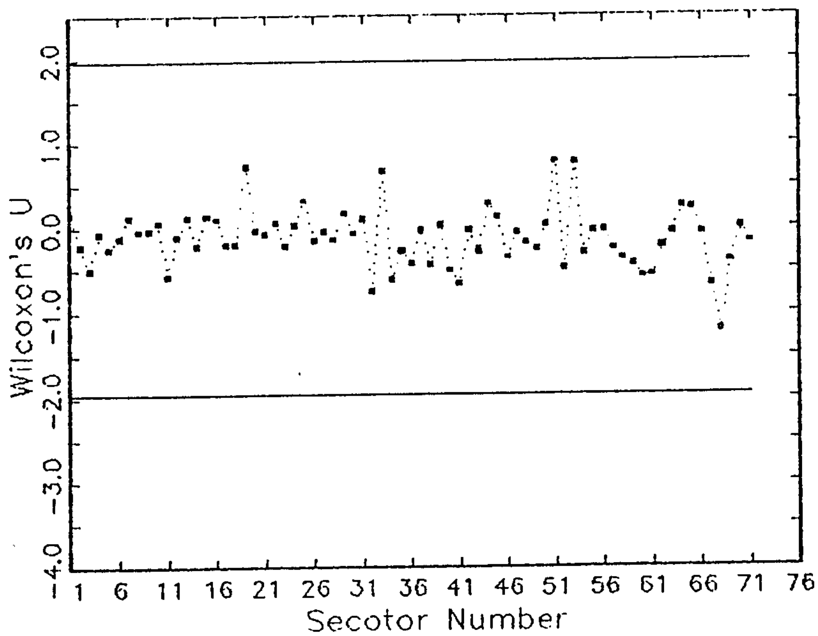


FIGURE A-27

WILCOXON NON-PARAMETRIC TEST FOR DIRECT MDLAGBIG

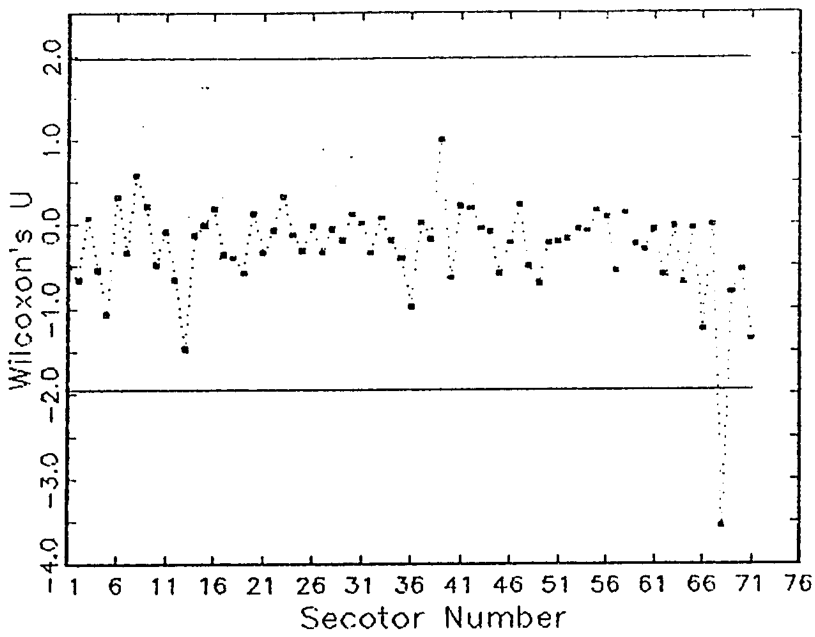


FIGURE A-28

WILCOXON NON-PARAMETRIC TEST FOR INVERSE MDLAGBIG

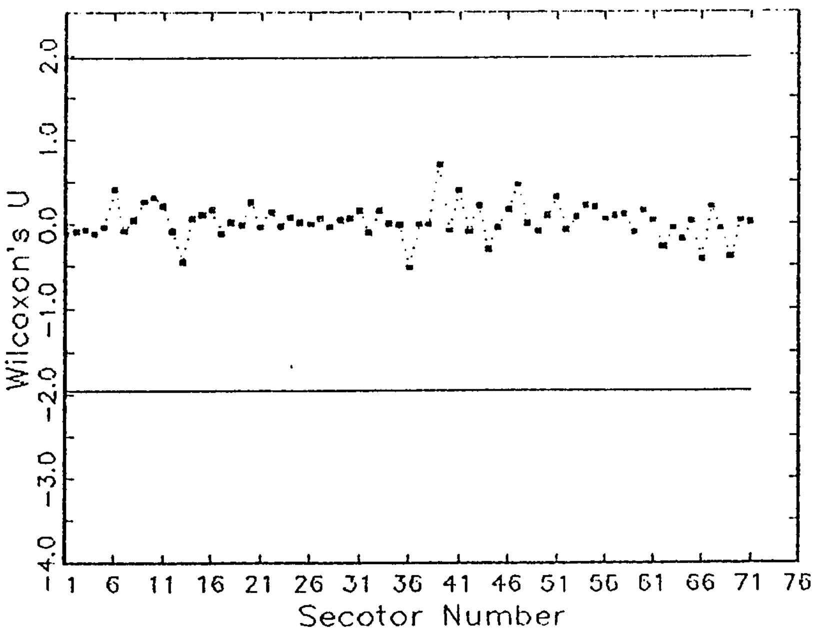


FIGURE A-29

WILCOXON NON-PARAMETRIC TEST FOR DIRECT MDLAGMIP

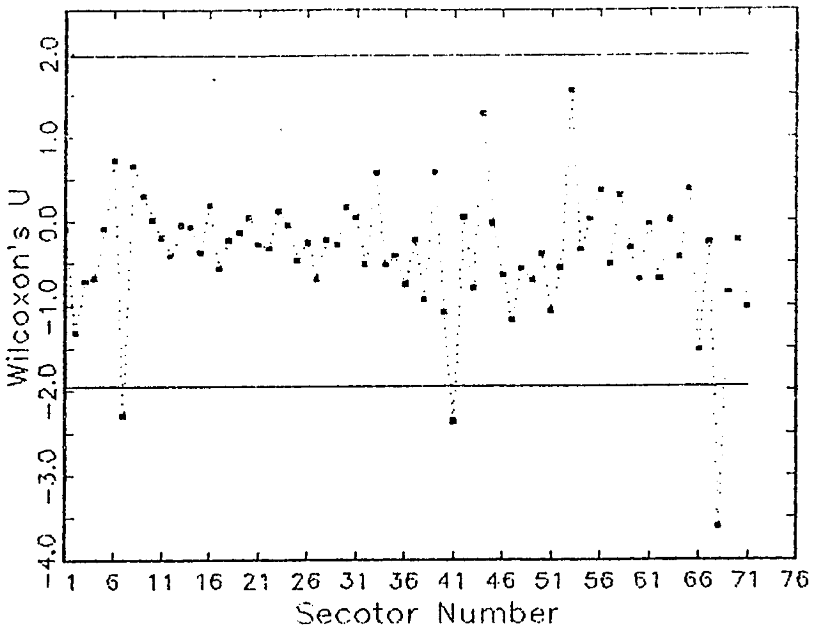


FIGURE A-30

WILCOXON NON-PARAMETRIC TEST FOR INVERSE MDLAGMIP

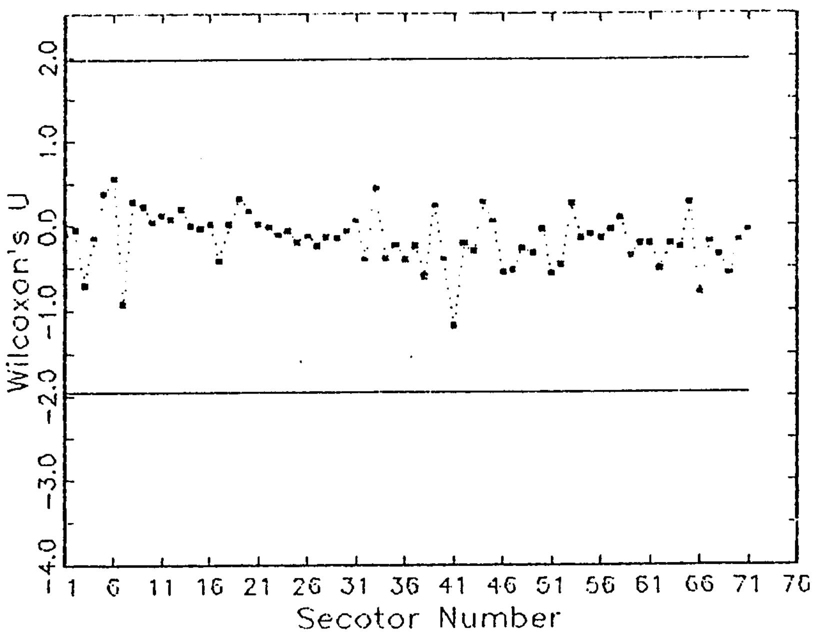


FIGURE A-31

WILCOXON NON-PARAMETRIC TEST FOR DIRECT MDRASKEY

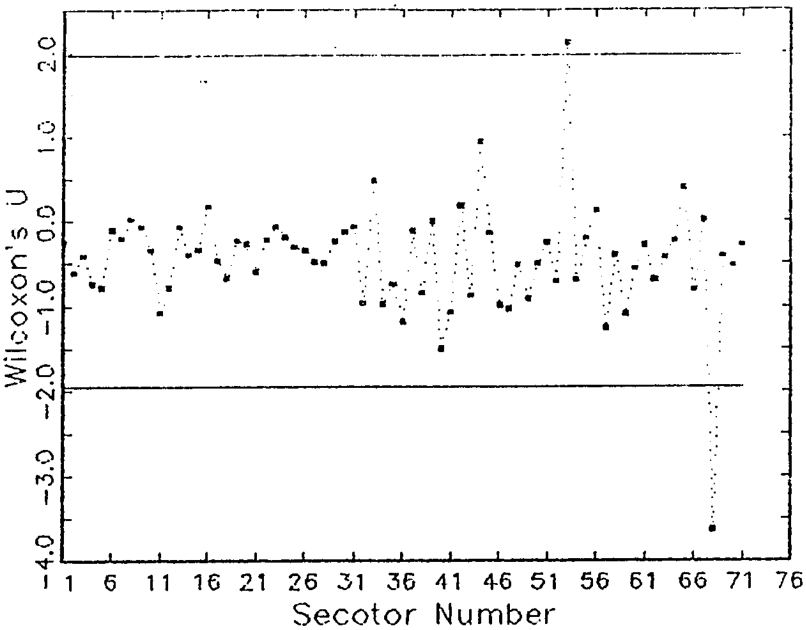


FIGURE A-32

WILCOXON NON-PARAMETRIC TEST FOR INVERSE MDRASKEY

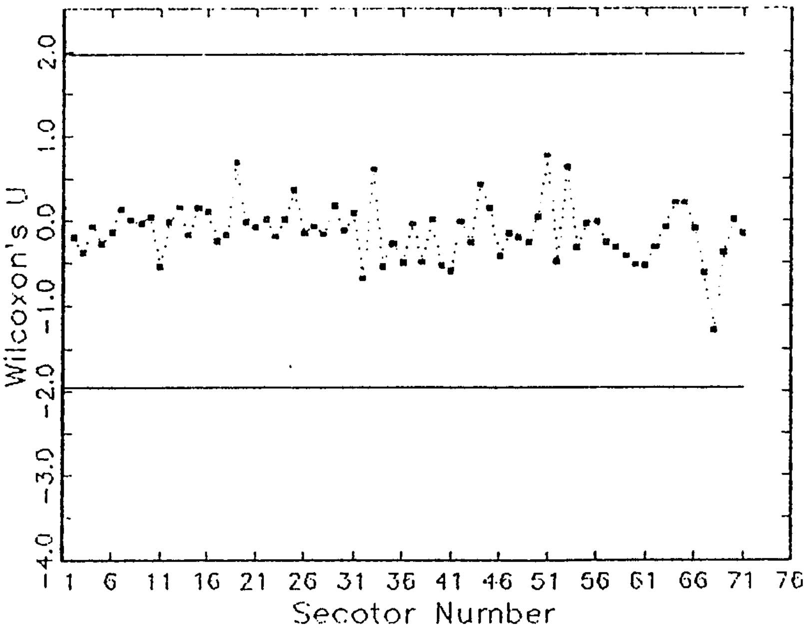


FIGURE A-33

WILCOXON NON-PARAMETRIC TEST FOR DIRECT MDRASBIG

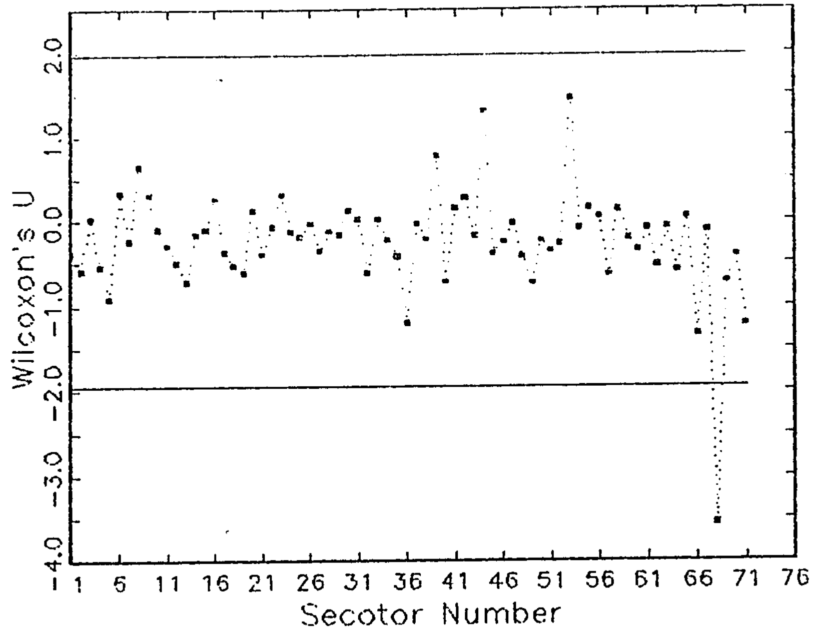


FIGURE A-34

WILCOXON NON-PARAMETRIC TEST FOR INVERSE MDRASBIG

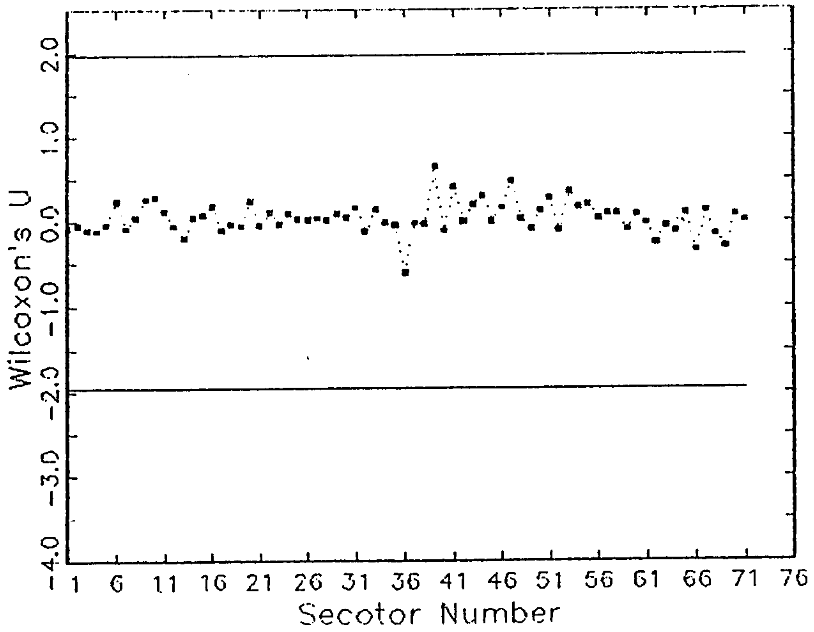


FIGURE A-35

WILCOXON NON-PARAMETRIC TEST FOR DIRECT MDRASMIP

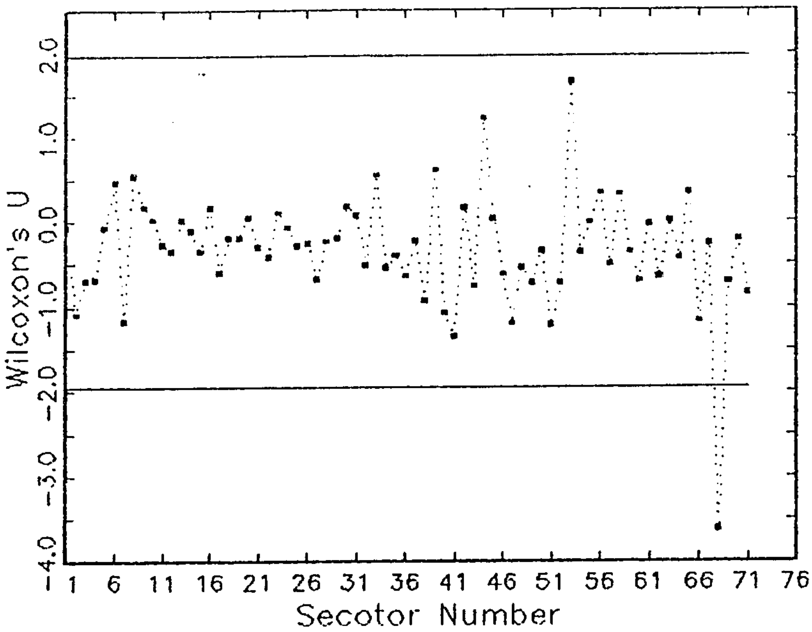


FIGURE A-36

WILCOXON NON-PARAMETRIC TEST FOR INVERSE MDRASMIP

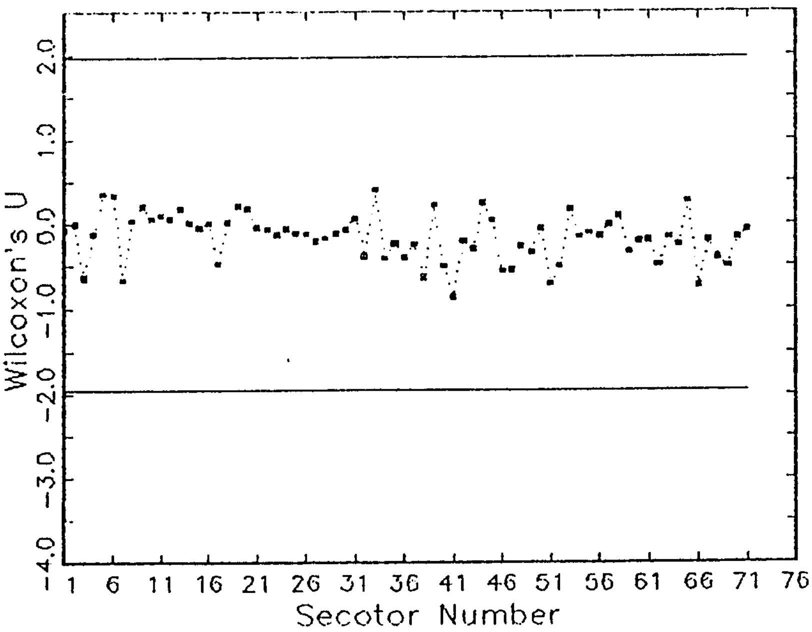


FIGURE A-37

WILCOXON NON-PARAMETRIC TEST FOR DIRECT RESMIN

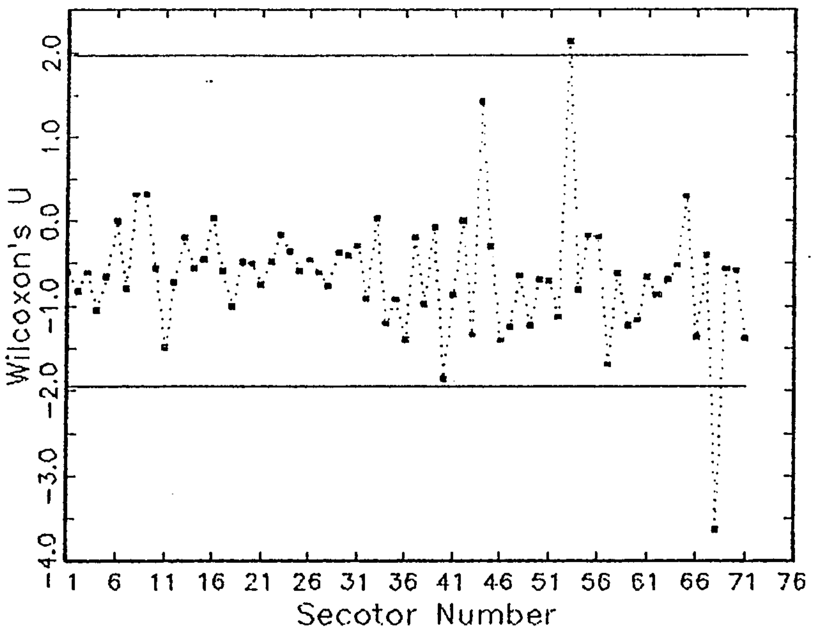
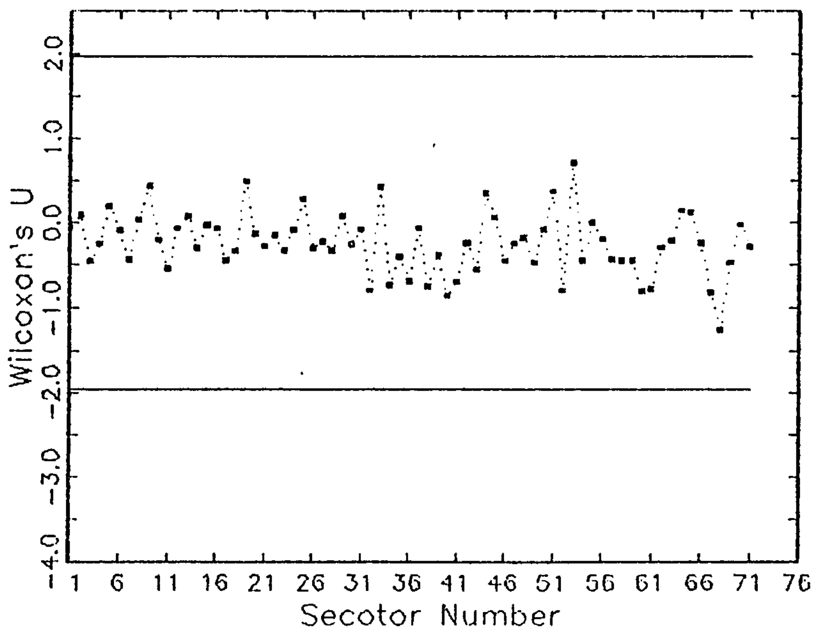


FIGURE A-38

WILCOXON NON-PARAMETRIC TEST FOR INVERSE RESMIN



APPENDIX B

FREQUENCY DISTRIBUTION OF COEFFICIENTS OF EQUALITY

FOR

ORIGINAL AND MODIFIED ESTIMATES

DIRECT AND INVERSE COEFFICIENTS

TABLE B-1

DISTRIBUTION OF C.O.E, ESTIMATED DIRECT NAIVE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	1131	21	2.1000	48
2	0.2000	150	22	2.2000	4
3	0.3000	144	23	2.3000	34
4	0.4000	196	24	2.4000	26
5	0.5000	213	25	2.5000	28
6	0.6000	234	26	2.6000	25
7	0.7000	227	27	2.7000	24
8	0.8000	275	28	2.8000	23
9	0.9000	323	29	2.9000	18
10	1.0000	278	30	3.0000	18
11	1.1000	248	31	3.1000	14
12	1.2000	212	32	3.2000	12
13	1.3000	161	33	3.3000	9
14	1.4000	149	34	3.4000	4
15	1.5000	98	35	3.5000	14
16	1.6000	91	36	3.6000	9
17	1.7000	80	37	3.7000	8
18	1.8000	76	38	3.8000	8
19	1.9000	68	39	3.9000	12
20	2.0000	43	40	1000.0	268

TABLE B-2

DISTRIBUTION OF C.O.E, ESTIMATED INVERSE NAIVE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	15	21	2.1000	41
2	0.2000	45	22	2.2000	42
3	0.3000	114	23	2.3000	17
4	0.4000	230	24	2.4000	10
5	0.5000	297	25	2.5000	17
6	0.6000	369	26	2.6000	16
7	0.7000	478	27	2.7000	9
8	0.8000	494	28	2.8000	3
9	0.9000	515	29	2.9000	12
10	1.0000	556	30	3.0000	6
11	1.1000	434	31	3.1000	6
12	1.2000	332	32	3.2000	7
13	1.3000	239	33	3.3000	3
14	1.4000	196	34	3.4000	1
15	1.5000	127	35	3.5000	2
16	1.6000	92	36	3.6000	10
17	1.7000	91	37	3.7000	1
18	1.8000	68	38	3.8000	3
19	1.9000	59	39	3.9000	5
20	2.0000	30	40	1000.0	49

TABLE B-3

DISTRIBUTION OF C.O.E, ESTIMATED DIRECT RAS COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	1101	21	2.1000	37
2	0.2000	141	22	2.2000	34
3	0.3000	159	23	2.3000	39
4	0.4000	170	24	2.4000	27
5	0.5000	194	25	2.5000	27
6	0.6000	216	26	2.6000	15
7	0.7000	254	27	2.7000	26
8	0.8000	269	28	2.8000	15
9	0.9000	317	29	2.9000	9
10	1.0000	305	30	3.0000	13
11	1.1000	260	31	3.1000	10
12	1.2000	243	32	3.2000	15
13	1.3000	179	33	3.3000	8
14	1.4000	143	34	3.4000	8
15	1.5000	120	35	3.5000	7
16	1.6000	114	36	3.6000	9
17	1.7000	76	37	3.7000	13
18	1.8000	73	38	3.8000	3
19	1.9000	68	39	3.9000	6
20	2.0000	59	40	1000.0	259

TABLE B-4

DISTRIBUTION OF C.O.E, ESTIMATED INVERSE RAS COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	20	21	2.1000	25
2	0.2000	53	22	2.2000	25
3	0.3000	45	23	2.3000	15
4	0.4000	73	24	2.4000	12
5	0.5000	128	25	2.5000	9
6	0.6000	209	26	2.6000	8
7	0.7000	319	27	2.7000	9
8	0.8000	484	28	2.8000	9
9	0.9000	626	29	2.9000	9
10	1.0000	738	30	3.0000	7
11	1.1000	651	31	3.1000	9
12	1.2000	497	32	3.2000	3
13	1.3000	333	33	3.3000	5
14	1.4000	233	34	3.4000	4
15	1.5000	134	35	3.5000	3
16	1.6000	110	36	3.6000	4
17	1.7000	71	37	3.7000	0
18	1.8000	48	38	3.8000	4
19	1.9000	36	39	3.9000	3
20	2.0000	31	40	1000.0	39

TABLE B-5

DISTRIBUTION OF C.O.E, ESTIMATED DIRECT RECRAS COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	1099	21	2.1000	38
2	0.2000	164	22	2.2000	38
3	0.3000	151	23	2.3000	28
4	0.4000	195	24	2.4000	33
5	0.5000	210	25	2.5000	25
6	0.6000	229	26	2.6000	23
7	0.7000	274	27	2.7000	13
8	0.8000	291	28	2.8000	13
9	0.9000	313	29	2.9000	13
10	1.0000	291	30	3.0000	10
11	1.1000	221	31	3.1000	13
12	1.2000	232	32	3.2000	8
13	1.3000	175	33	3.3000	17
14	1.4000	147	34	3.4000	11
15	1.5000	107	35	3.5000	4
16	1.6000	102	36	3.6000	14
17	1.7000	70	37	3.7000	6
18	1.8000	70	38	3.8000	7
19	1.9000	66	39	3.9000	11
20	2.0000	53	40	1000.0	256

TABLE B-6

DISTRIBUTION OF C.O.E, ESTIMATED INVERSE RECRAS COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	20	21	2.1000	28
2	0.2000	60	22	2.2000	20
3	0.3000	58	23	2.3000	23
4	0.4000	115	24	2.4000	18
5	0.5000	214	25	2.5000	13
6	0.6000	356	26	2.6000	13
7	0.7000	451	27	2.7000	12
8	0.8000	537	28	2.8000	11
9	0.9000	524	29	2.9000	10
10	1.0000	564	30	3.0000	6
11	1.1000	543	31	3.1000	5
12	1.2000	375	32	3.2000	6
13	1.3000	272	33	3.3000	6
14	1.4000	214	34	3.4000	3
15	1.5000	154	35	3.5000	5
16	1.6000	100	36	3.6000	2
17	1.7000	83	37	3.7000	7
18	1.8000	69	38	3.8000	2
19	1.9000	50	39	3.9000	2
20	2.0000	44	40	1000.0	46

TABLE B-7

DISTRIBUTION OF C.O.E, ESTIMATED DIRECT PROPVA COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	1142	21	2.1000	50
2	0.2000	162	22	2.2000	44
3	0.3000	164	23	2.3000	31
4	0.4000	238	24	2.4000	20
5	0.5000	230	25	2.5000	22
6	0.6000	256	26	2.6000	27
7	0.7000	290	27	2.7000	16
8	0.8000	288	28	2.8000	10
9	0.9000	296	29	2.9000	13
10	1.0000	310	30	3.0000	12
11	1.1000	241	31	3.1000	6
12	1.2000	202	32	3.2000	12
13	1.3000	150	33	3.3000	16
14	1.4000	114	34	3.4000	7
15	1.5000	101	35	3.5000	10
16	1.6000	85	36	3.6000	7
17	1.7000	64	37	3.7000	10
18	1.8000	52	38	3.8000	7
19	1.9000	53	39	3.9000	3
20	2.0000	40	40	1000.0	240

TABLE B-8

DISTRIBUTION OF C.O.E, ESTIMATED INVERSE PROPVA COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	25	21	2.1000	32
2	0.2000	108	22	2.2000	16
3	0.3000	217	23	2.3000	16
4	0.4000	327	24	2.4000	12
5	0.5000	494	25	2.5000	14
6	0.6000	597	26	2.6000	4
7	0.7000	542	27	2.7000	6
8	0.8000	542	28	2.8000	7
9	0.9000	460	29	2.9000	6
10	1.0000	451	30	3.0000	8
11	1.1000	308	31	3.1000	8
12	1.2000	224	32	3.2000	5
13	1.3000	153	33	3.3000	6
14	1.4000	120	34	3.4000	3
15	1.5000	69	35	3.5000	3
16	1.6000	63	36	3.6000	5
17	1.7000	39	37	3.7000	1
18	1.8000	49	38	3.8000	4
19	1.9000	39	39	3.9000	1
20	2.0000	27	40	1000.0	30

TABLE B-9

DISTRIBUTION OF C.O.E, ESTIMATED DIRECT FRIED COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	1112	21	2.1000	44
2	0.2000	138	22	2.2000	34
3	0.3000	165	23	2.3000	42
4	0.4000	187	24	2.4000	26
5	0.5000	192	25	2.5000	20
6	0.6000	234	26	2.6000	23
7	0.7000	228	27	2.7000	14
8	0.8000	270	28	2.8000	18
9	0.9000	317	29	2.9000	11
10	1.0000	281	30	3.0000	7
11	1.1000	256	31	3.1000	15
12	1.2000	227	32	3.2000	14
13	1.3000	193	33	3.3000	13
14	1.4000	153	34	3.4000	7
15	1.5000	116	35	3.5000	5
16	1.6000	106	36	3.6000	17
17	1.7000	89	37	3.7000	8
18	1.8000	74	38	3.8000	8
19	1.9000	53	39	3.9000	5
20	2.0000	63	40	1000.0	256

TABLE B-10

DISTRIBUTION OF C.O.E, ESTIMATED INVERSE FRIED COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	16	21	2.1000	21
2	0.20000	50	22	2.2000	26
3	0.30000	45	23	2.3000	23
4	0.40000	86	24	2.4000	14
5	0.50000	117	25	2.5000	9
6	0.60000	217	26	2.6000	6
7	0.70000	320	27	2.7000	9
8	0.80000	474	28	2.8000	7
9	0.90000	634	29	2.9000	11
10	1.00000	723	30	3.0000	15
11	1.10000	644	31	3.1000	4
12	1.20000	462	32	3.2000	5
13	1.30000	345	33	3.3000	2
14	1.40000	242	34	3.4000	4
15	1.50000	142	35	3.5000	2
16	1.60000	122	36	3.6000	6
17	1.70000	75	37	3.7000	2
18	1.80000	46	38	3.8000	5
19	1.90000	36	39	3.9000	1
20	2.00000	34	40	1000.0	39

TABLE B-11

DISTRIBUTION OF C.O.E, ESTIMATED DIRECT RECLAG COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	1227	21	2.1000	22
2	0.2000	285	22	2.2000	16
3	0.3000	332	23	2.3000	12
4	0.4000	328	24	2.4000	21
5	0.5000	378	25	2.5000	11
6	0.6000	436	26	2.6000	11
7	0.7000	381	27	2.7000	12
8	0.8000	308	28	2.8000	10
9	0.9000	213	29	2.9000	10
10	1.0000	173	30	3.0000	4
11	1.1000	137	31	3.1000	5
12	1.2000	106	32	3.2000	11
13	1.3000	90	33	3.3000	6
14	1.4000	79	34	3.4000	8
15	1.5000	47	35	3.5000	2
16	1.6000	45	36	3.6000	9
17	1.7000	46	37	3.7000	3
18	1.8000	37	38	3.8000	4
19	1.9000	25	39	3.9000	5
20	2.0000	23	40	1000.0	163

TABLE B-12

DISTRIBUTION OF C.O.E, ESTIMATED INVERSE RECLAG COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	40	21	2.1000	17
2	0.2000	108	22	2.2000	11
3	0.3000	228	23	2.3000	5
4	0.4000	428	24	2.4000	5
5	0.5000	618	25	2.5000	5
6	0.6000	709	26	2.6000	5
7	0.7000	643	27	2.7000	14
8	0.8000	538	28	2.8000	4
9	0.9000	445	29	2.9000	2
10	1.0000	325	30	3.0000	3
11	1.1000	256	31	3.1000	1
12	1.2000	186	32	3.2000	4
13	1.3000	110	33	3.3000	4
14	1.4000	93	34	3.4000	1
15	1.5000	64	35	3.5000	0
16	1.6000	37	36	3.6000	1
17	1.7000	33	37	3.7000	0
18	1.8000	28	38	3.8000	- 1
19	1.9000	27	39	3.9000	1
20	2.0000	17	40	1000.0	24

TABLE B-13

DISTRIBUTION OF C.O.E, ESTIMATED DIRECT RASLAG COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	1101	21	2.1000	37
2	0.2000	141	22	2.2000	34
3	0.3000	159	23	2.3000	39
4	0.4000	170	24	2.4000	27
5	0.5000	194	25	2.5000	27
6	0.6000	216	26	2.6000	15
7	0.7000	254	27	2.7000	26
8	0.8000	269	28	2.8000	15
9	0.9000	317	29	2.9000	9
10	1.0000	305	30	3.0000	13
11	1.1000	260	31	3.1000	10
12	1.2000	243	32	3.2000	15
13	1.3000	179	33	3.3000	8
14	1.4000	143	34	3.4000	8
15	1.5000	120	35	3.5000	7
16	1.6000	114	36	3.6000	9
17	1.7000	76	37	3.7000	13
18	1.8000	73	38	3.8000	3
19	1.9000	68	39	3.9000	6
20	2.0000	59	40	1000.0	259

TABLE B-14

DISTRIBUTION OF C.O.E, ESTIMATED INVERSE RASLAG COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	20	21	2.1000	25
2	0.2000	53	22	2.2000	25
3	0.3000	45	23	2.3000	15
4	0.4000	73	24	2.4000	12
5	0.5000	128	25	2.5000	9
6	0.6000	209	26	2.6000	8
7	0.7000	319	27	2.7000	9
8	0.8000	484	28	2.8000	9
9	0.9000	626	29	2.9000	9
10	1.0000	738	30	3.0000	7
11	1.1000	651	31	3.1000	9
12	1.2000	497	32	3.2000	3
13	1.3000	333	33	3.3000	5
14	1.4000	233	34	3.4000	4
15	1.5000	134	35	3.5000	3
16	1.6000	110	36	3.6000	4
17	1.7000	71	37	3.7000	0
18	1.8000	48	38	3.8000	- 4
19	1.9000	36	39	3.9000	3
20	2.0000	31	40	1000.0	39

TABLE B-15

DISTRIBUTION OF C.O.E, ESTIMATED DIRECT RERALA COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	1099	21	2.1000	37
2	0.2000	165	22	2.2000	39
3	0.3000	149	23	2.3000	25
4	0.4000	197	24	2.4000	36
5	0.5000	207	25	2.5000	24
6	0.6000	231	26	2.6000	21
7	0.7000	273	27	2.7000	14
8	0.8000	293	28	2.8000	14
9	0.9000	307	29	2.9000	12
10	1.0000	292	30	3.0000	11
11	1.1000	228	31	3.1000	13
12	1.2000	226	32	3.2000	7
13	1.3000	176	33	3.3000	18
14	1.4000	150	34	3.4000	11
15	1.5000	103	35	3.5000	4
16	1.6000	103	36	3.6000	12
17	1.7000	72	37	3.7000	7
18	1.8000	70	38	3.8000	8
19	1.9000	65	39	3.9000	11
20	2.0000	55	40	1000.0	256

TABLE B-16

DISTRIBUTION OF C.O.E, ESTIMATED INVERSE RERALA COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	20	21	2.1000	36
2	0.2000	56	22	2.2000	19
3	0.3000	56	23	2.3000	19
4	0.4000	90	24	2.4000	17
5	0.5000	166	25	2.5000	16
6	0.6000	279	26	2.6000	11
7	0.7000	384	27	2.7000	13
8	0.8000	504	28	2.8000	7
9	0.9000	609	29	2.9000	9
10	1.0000	636	30	3.0000	8
11	1.1000	567	31	3.1000	7
12	1.2000	388	32	3.2000	4
13	1.3000	321	33	3.3000	7
14	1.4000	223	34	3.4000	6
15	1.5000	159	35	3.5000	3
16	1.6000	98	36	3.6000	2
17	1.7000	91	37	3.7000	5
18	1.8000	60	38	3.8000	2
19	1.9000	56	39	3.9000	4
20	2.0000	38	40	1000.0	45

TABLE B-17

DISTRIBUTION OF C.O.E, ESTIMATED ALMON DIRECT COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	2306	21	2.1000	21
2	0.2000	39	22	2.2000	26
3	0.3000	43	23	2.3000	27
4	0.4000	75	24	2.4000	25
5	0.5000	86	25	2.5000	25
6	0.6000	119	26	2.6000	16
7	0.7000	105	27	2.7000	23
8	0.8000	101	28	2.8000	12
9	0.9000	104	29	2.9000	17
10	1.0000	85	30	3.0000	14
11	1.1000	74	31	3.1000	15
12	1.2000	55	32	3.2000	19
13	1.3000	54	33	3.3000	13
14	1.4000	42	34	3.4000	9
15	1.5000	51	35	3.5000	8
16	1.6000	43	36	3.6000	8
17	1.7000	24	37	3.7000	20
18	1.8000	32	38	3.8000	13
19	1.9000	28	39	3.9000	13
20	2.0000	36	40	4000.0	1014

TABLE B-18

DISTRIBUTION OF C.O.E, ESTIMATED INVERSE ALMON COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.1000	1800	21	2.1000	38
2	0.2000	119	22	2.2000	26
3	0.3000	138	23	2.3000	25
4	0.4000	191	24	2.4000	16
5	0.5000	274	25	2.5000	23
6	0.6000	290	26	2.6000	13
7	0.7000	294	27	2.7000	19
8	0.8000	226	28	2.8000	21
9	0.9000	186	29	2.9000	13
10	1.0000	213	30	3.0000	20
11	1.1000	98	31	3.1000	13
12	1.2000	89	32	3.2000	10
13	1.3000	75	33	3.3000	14
14	1.4000	55	34	3.4000	10
15	1.5000	55	35	3.5000	6
16	1.6000	43	36	3.6000	15
17	1.7000	51	37	3.7000	7
18	1.8000	38	38	3.8000	9
19	1.9000	27	39	3.9000	17
20	2.0000	28	40	1000.0	436

TABLE B-19

DISTRIBUTION OF C.O.E ESTIMATED DIRECT MDNAVKEY COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	1068	21	2.1000	45
2	0.20000	145	22	2.2000	37
3	0.30000	126	23	2.3000	30
4	0.40000	162	24	2.4000	24
5	0.50000	178	25	2.5000	25
6	0.60000	204	26	2.6000	22
7	0.70000	203	27	2.7000	22
8	0.80000	254	28	2.8000	22
9	0.90000	294	29	2.9000	17
10	1.00000	701	30	3.0000	17
11	1.10000	231	31	3.1000	13
12	1.20000	198	32	3.2000	11
13	1.30000	155	33	3.3000	8
14	1.40000	138	34	3.4000	4
15	1.50000	89	35	3.5000	13
16	1.60000	87	36	3.6000	8
17	1.70000	73	37	3.7000	8
18	1.80000	70	38	3.8000	6
19	1.90000	62	39	3.9000	10
20	2.00000	38	40	1000.0	223

TABLE B-20

DISTRIBUTION OF C.O.E ESTIMATED INVERSE MDNAVKEY COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	1131	21	2.1000	44
2	0.20000	150	22	2.2000	36
3	0.30000	142	23	2.3000	34
4	0.40000	189	24	2.4000	26
5	0.50000	208	25	2.5000	27
6	0.60000	220	26	2.6000	22
7	0.70000	209	27	2.7000	23
8	0.80000	233	28	2.8000	20
9	0.90000	257	29	2.9000	17
10	1.00000	708	30	3.0000	14
11	1.10000	187	31	3.1000	12
12	1.20000	174	32	3.2000	10
13	1.30000	127	33	3.3000	9
14	1.40000	127	34	3.4000	4
15	1.50000	85	35	3.5000	13
16	1.60000	72	36	3.6000	9
17	1.70000	72	37	3.7000	8
18	1.80000	67	38	3.8000	7
19	1.90000	59	39	3.9000	11
20	2.00000	39	40	1000.0	238

TABLE B-21

DISTRIBUTION OF C.O.E ESTIMATED DIRECT MDNAVBIG COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	1129	21	2.1000	39
2	0.20000	147	22	2.2000	35
3	0.30000	135	23	2.3000	32
4	0.40000	186	24	2.4000	21
5	0.50000	201	25	2.5000	28
6	0.60000	207	26	2.6000	23
7	0.70000	210	27	2.7000	22
8	0.80000	239	28	2.8000	18
9	0.90000	275	29	2.9000	14
10	1.00000	736	30	3.0000	16
11	1.10000	194	31	3.1000	12
12	1.20000	178	32	3.2000	11
13	1.30000	135	33	3.3000	9
14	1.40000	129	34	3.4000	4
15	1.50000	85	35	3.5000	14
16	1.60000	71	36	3.6000	8
17	1.70000	74	37	3.7000	8
18	1.80000	58	38	3.8000	7
19	1.90000	60	39	3.9000	10
20	2.00000	37	40	1000.0	223

TABLE B-22

DISTRIBUTION OF C.O.E ESTIMATED INVERSE MDNAVBIG COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	1051	21	2.1000	38
2	0.20000	130	22	2.2000	26
3	0.30000	160	23	2.3000	35
4	0.40000	185	24	2.4000	31
5	0.50000	180	25	2.5000	18
6	0.60000	212	26	2.6000	15
7	0.70000	211	27	2.7000	19
8	0.80000	234	28	2.8000	12
9	0.90000	303	29	2.9000	16
10	1.00000	693	30	3.0000	10
11	1.10000	243	31	3.1000	13
12	1.20000	193	32	3.2000	11
13	1.30000	162	33	3.3000	5
14	1.40000	134	34	3.4000	9
15	1.50000	108	35	3.5000	6
16	1.60000	97	36	3.6000	11
17	1.70000	71	37	3.7000	6
18	1.80000	66	38	3.8000	6
19	1.90000	55	39	3.9000	7
20	2.00000	48	40	1000.0	211

TABLE B-23

DISTRIBUTION OF C.O.E ESTIMATED DIRECT MDNAVMIP COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	1097	21	2.1000	49
2	0.20000	112	22	2.2000	34
3	0.30000	125	23	2.3000	33
4	0.40000	159	24	2.4000	34
5	0.50000	156	25	2.5000	21
6	0.60000	179	26	2.6000	24
7	0.70000	186	27	2.7000	23
8	0.80000	227	28	2.8000	23
9	0.90000	193	29	2.9000	20
10	1.00000	707	30	3.0000	13
11	1.10000	209	31	3.1000	18
12	1.20000	176	32	3.2000	15
13	1.30000	187	33	3.3000	19
14	1.40000	159	34	3.4000	12
15	1.50000	127	35	3.5000	14
16	1.60000	113	36	3.6000	9
17	1.70000	83	37	3.7000	9
18	1.80000	72	38	3.8000	10
19	1.90000	72	39	3.9000	8
20	2.00000	57	40	1000.0	257

TABLE B-24

DISTRIBUTION OF C.O.E ESTIMATED INVERSE MDNAVMIP COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	1088	21	2.1000	37
2	0.20000	124	22	2.2000	48
3	0.30000	142	23	2.3000	35
4	0.40000	146	24	2.4000	27
5	0.50000	180	25	2.5000	34
6	0.60000	168	26	2.6000	28
7	0.70000	193	27	2.7000	26
8	0.80000	208	28	2.8000	23
9	0.90000	212	29	2.9000	13
10	1.00000	718	30	3.0000	14
11	1.10000	231	31	3.1000	16
12	1.20000	223	32	3.2000	12
13	1.30000	169	33	3.3000	13
14	1.40000	162	34	3.4000	8
15	1.50000	109	35	3.5000	7
16	1.60000	78	36	3.6000	9
17	1.70000	66	37	3.7000	9
18	1.80000	68	38	3.8000	10
19	1.90000	59	39	3.9000	6
20	2.00000	55	40	1000.0	267

TABLE B-25

DISTRIBUTION OF C.O.E ESTIMATED DIRECT MDLAGKEY COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	1041	21	2.1000	33
2	0.20000	132	22	2.2000	32
3	0.30000	145	23	2.3000	31
4	0.40000	176	24	2.4000	27
5	0.50000	177	25	2.5000	23
6	0.60000	199	26	2.6000	20
7	0.70000	234	27	2.7000	19
8	0.80000	231	28	2.8000	14
9	0.90000	308	29	2.9000	13
10	1.00000	728	30	3.0000	11
11	1.10000	223	31	3.1000	9
12	1.20000	208	32	3.2000	11
13	1.30000	166	33	3.3000	10
14	1.40000	114	34	3.4000	11
15	1.50000	116	35	3.5000	4
16	1.60000	87	36	3.6000	7
17	1.70000	80	37	3.7000	7
18	1.80000	64	38	3.8000	6
19	1.90000	64	39	3.9000	2
20	2.00000	42	40	1000.0	216

TABLE B-26

DISTRIBUTION OF C.O.E ESTIMATED INVERSE MDLAGKEY COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	1077	21	2.1000	46
2	0.20000	105	22	2.2000	32
3	0.30000	115	23	2.3000	36
4	0.40000	155	24	2.4000	34
5	0.50000	173	25	2.5000	38
6	0.60000	166	26	2.6000	25
7	0.70000	208	27	2.7000	25
8	0.80000	219	28	2.8000	22
9	0.90000	202	29	2.9000	19
10	1.00000	718	30	3.0000	19
11	1.10000	213	31	3.1000	16
12	1.20000	205	32	3.2000	17
13	1.30000	173	33	3.3000	12
14	1.40000	157	34	3.4000	13
15	1.50000	119	35	3.5000	7
16	1.60000	101	36	3.6000	11
17	1.70000	80	37	3.7000	13
18	1.80000	68	38	3.8000	6
19	1.90000	77	39	3.9000	7
20	2.00000	50	40	1000.0	262

TABLE B-27

DISTRIBUTION OF C.O.E ESTIMATED DIRECT MDLAGBIG COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	1073	21	2.1000	40
2	0.20000	110	22	2.2000	35
3	0.30000	141	23	2.3000	33
4	0.40000	148	24	2.4000	33
5	0.50000	171	25	2.5000	24
6	0.60000	170	26	2.6000	34
7	0.70000	218	27	2.7000	24
8	0.80000	208	28	2.8000	14
9	0.90000	220	29	2.9000	20
10	1.00000	728	30	3.0000	11
11	1.10000	242	31	3.1000	13
12	1.20000	214	32	3.2000	15
13	1.30000	181	33	3.3000	14
14	1.40000	147	34	3.4000	7
15	1.50000	101	35	3.5000	14
16	1.60000	80	36	3.6000	5
17	1.70000	72	37	3.7000	14
18	1.80000	80	38	3.8000	9
19	1.90000	60	39	3.9000	3
20	2.00000	47	40	1000.0	268

TABLE B-28

DISTRIBUTION OF C.O.E ESTIMATED INVERSE MDLAGBIG COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	1138	21	2.1000	29
2	0.20000	164	22	2.2000	28
3	0.30000	190	23	2.3000	33
4	0.40000	207	24	2.4000	27
5	0.50000	216	25	2.5000	17
6	0.60000	233	26	2.6000	21
7	0.70000	282	27	2.7000	14
8	0.80000	320	28	2.8000	12
9	0.90000	286	29	2.9000	15
10	1.00000	321	30	3.0000	10
11	1.10000	252	31	3.1000	8
12	1.20000	225	32	3.2000	6
13	1.30000	160	33	3.3000	9
14	1.40000	119	34	3.4000	11
15	1.50000	104	35	3.5000	7
16	1.60000	85	36	3.6000	11
17	1.70000	71	37	3.7000	8
18	1.80000	65	38	3.8000	5
19	1.90000	55	39	3.9000	12
20	2.00000	39	40	1000.0	226

TABLE B-29

DISTRIBUTION OF C.O.E ESTIMATED MDLAGMIP DIRECT COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	14	21	2.1000	30
2	0.20000	43	22	2.2000	29
3	0.30000	75	23	2.3000	29
4	0.40000	152	24	2.4000	11
5	0.50000	252	25	2.5000	17
6	0.60000	326	26	2.6000	11
7	0.70000	422	27	2.7000	6
8	0.80000	506	28	2.8000	4
9	0.90000	542	29	2.9000	11
10	1.00000	697	30	3.0000	7
11	1.10000	586	31	3.1000	7
12	1.20000	337	32	3.2000	6
13	1.30000	225	33	3.3000	4
14	1.40000	188	34	3.4000	1
15	1.50000	121	35	3.5000	1
16	1.60000	109	36	3.6000	4
17	1.70000	75	37	3.7000	4
18	1.80000	54	38	3.8000	2
19	1.90000	52	39	3.9000	4
20	2.00000	31	40	1000.0	46

TABLE B-30

DISTRIBUTION OF C.O.E ESTIMATED INVERSE MDLAGMIP COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	16	21	2.1000	10
2	0.20000	46	22	2.2000	18
3	0.30000	77	23	2.3000	10
4	0.40000	138	24	2.4000	8
5	0.50000	259	25	2.5000	6
6	0.60000	342	26	2.6000	5
7	0.70000	417	27	2.7000	2
8	0.80000	572	28	2.8000	5
9	0.90000	680	29	2.9000	2
10	1.00000	1027	30	3.0000	5
11	1.10000	523	31	3.1000	4
12	1.20000	256	32	3.2000	2
13	1.30000	177	33	3.3000	0
14	1.40000	113	34	3.4000	2
15	1.50000	85	35	3.5000	2
16	1.60000	66	36	3.6000	2
17	1.70000	51	37	3.7000	1
18	1.80000	31	38	3.8000	2
19	1.90000	31	39	3.9000	1
20	2.00000	25	40	1000.0	22

TABLE B-31

DISTRIBUTION OF C.O.E ESTIMATED DIRECT MDRASKEY COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	14	21	2.1000	13
2	0.20000	42	22	2.2000	7
3	0.30000	77	23	2.3000	5
4	0.40000	122	24	2.4000	5
5	0.50000	250	25	2.5000	3
6	0.60000	349	26	2.6000	4
7	0.70000	455	27	2.7000	2
8	0.80000	677	28	2.8000	5
9	0.90000	856	29	2.9000	1
10	1.00000	997	30	3.0000	1
11	1.10000	475	31	3.1000	4
12	1.20000	222	32	3.2000	1
13	1.30000	141	33	3.3000	4
14	1.40000	92	34	3.4000	1
15	1.50000	57	35	3.5000	2
16	1.60000	43	36	3.6000	0
17	1.70000	41	37	3.7000	0
18	1.80000	34	38	3.8000	0
19	1.90000	16	39	3.9000	0
20	2.00000	17	40	1000.0	6

TABLE B-32

DISTRIBUTION OF C.O.E ESTIMATED INVERSE MDRASKEY COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	15	21	2.1000	23
2	0.20000	48	22	2.2000	24
3	0.30000	48	23	2.3000	16
4	0.40000	80	24	2.4000	13
5	0.50000	122	25	2.5000	13
6	0.60000	220	26	2.6000	4
7	0.70000	307	27	2.7000	8
8	0.80000	476	28	2.8000	3
9	0.90000	565	29	2.9000	8
10	1.00000	818	30	3.0000	15
11	1.10000	850	31	3.1000	5
12	1.20000	437	32	3.2000	3
13	1.30000	296	33	3.3000	2
14	1.40000	178	34	3.4000	4
15	1.50000	121	35	3.5000	4
16	1.60000	102	36	3.6000	5
17	1.70000	61	37	3.7000	2
18	1.80000	47	38	3.8000	3
19	1.90000	29	39	3.9000	3
20	2.00000	28	40	1000.0	35

TABLE B-33

DISTRIBUTION OF C.O.E ESTIMATED DIRECT MDRASBIG COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	17	21	2.1000	25
2	0.20000	38	22	2.2000	22
3	0.30000	44	23	2.3000	21
4	0.40000	58	24	2.4000	16
5	0.50000	75	25	2.5000	11
6	0.60000	126	26	2.6000	7
7	0.70000	215	27	2.7000	4
8	0.80000	350	28	2.8000	8
9	0.90000	487	29	2.9000	8
10	1.00000	1128	30	3.0000	3
11	1.10000	933	31	3.1000	5
12	1.20000	487	32	3.2000	6
13	1.30000	321	33	3.3000	5
14	1.40000	188	34	3.4000	3
15	1.50000	117	35	3.5000	2
16	1.60000	83	36	3.6000	1
17	1.70000	69	37	3.7000	3
18	1.80000	57	38	3.8000	2
19	1.90000	44	39	3.9000	0
20	2.00000	21	40	1000.0	31

TABLE B-34

DISTRIBUTION OF C.O.E ESTIMATED INVERSE MDRASBIG COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	4	21	2.1000	12
2	0.20000	19	22	2.2000	16
3	0.30000	32	23	2.3000	11
4	0.40000	61	24	2.4000	7
5	0.50000	99	25	2.5000	7
6	0.60000	161	26	2.6000	6
7	0.70000	255	27	2.7000	2
8	0.80000	384	28	2.8000	4
9	0.90000	574	29	2.9000	2
10	1.00000	1116	30	3.0000	3
11	1.10000	1043	31	3.1000	1
12	1.20000	510	32	3.2000	0
13	1.30000	275	33	3.3000	4
14	1.40000	161	34	3.4000	3
15	1.50000	85	35	3.5000	2
16	1.60000	61	36	3.6000	1
17	1.70000	45	37	3.7000	0
18	1.80000	28	38	3.8000	0
19	1.90000	24	39	3.9000	0
20	2.00000	13	40	1000.0	10

TABLE B-35

DISTRIBUTION OF C.O.E ESTIMATED DIRECT MDRASMIP COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	20	21	2.1000	28
2	0.20000	49	22	2.2000	13
3	0.30000	48	23	2.3000	19
4	0.40000	67	24	2.4000	14
5	0.50000	136	25	2.5000	6
6	0.60000	204	26	2.6000	9
7	0.70000	312	27	2.7000	6
8	0.80000	462	28	2.8000	7
9	0.90000	572	29	2.9000	7
10	1.00000	845	30	3.0000	7
11	1.10000	844	31	3.1000	6
12	1.20000	447	32	3.2000	5
13	1.30000	295	33	3.3000	1
14	1.40000	172	34	3.4000	6
15	1.50000	128	35	3.5000	2
16	1.60000	85	36	3.6000	4
17	1.70000	62	37	3.7000	4
18	1.80000	53	38	3.8000	1
19	1.90000	28	39	3.9000	5
20	2.00000	26	40	1000.0	36

TABLE B-36

DISTRIBUTION OF C.O.E ESTIMATED INVERSE MDRASMIP COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	26	21	2.1000	18
2	0.20000	45	22	2.2000	23
3	0.30000	37	23	2.3000	18
4	0.40000	43	24	2.4000	17
5	0.50000	75	25	2.5000	11
6	0.60000	133	26	2.6000	10
7	0.70000	211	27	2.7000	8
8	0.80000	342	28	2.8000	2
9	0.90000	475	29	2.9000	11
10	1.00000	1153	30	3.0000	3
11	1.10000	948	31	3.1000	5
12	1.20000	492	32	3.2000	8
13	1.30000	295	33	3.3000	6
14	1.40000	187	34	3.4000	2
15	1.50000	122	35	3.5000	2
16	1.60000	82	36	3.6000	1
17	1.70000	66	37	3.7000	2
18	1.80000	59	38	3.8000	0
19	1.90000	35	39	3.9000	2
20	2.00000	36	40	1000.0	30

TABLE B-37

DISTRIBUTION OF C.O E ESTIMATED DIRECT RESMIN COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	4	21	2.1000	9
2	0.20000	17	22	2.2000	10
3	0.30000	29	23	2.3000	13
4	0.40000	55	24	2.4000	5
5	0.50000	96	25	2.5000	6
6	0.60000	157	26	2.6000	5
7	0.70000	256	27	2.7000	2
8	0.80000	405	28	2.8000	5
9	0.90000	575	29	2.9000	5
10	1.00000	1133	30	3.0000	2
11	1.10000	1043	31	3.1000	2
12	1.20000	529	32	3.2000	1
13	1.30000	260	33	3.3000	2
14	1.40000	143	34	3.4000	0
15	1.50000	100	35	3.5000	0
16	1.60000	47	36	3.6000	1
17	1.70000	45	37	3.7000	3
18	1.80000	30	38	3.8000	1
19	1.90000	20	39	3.9000	0
20	2.00000	15	40	1000.0	10

TABLE B-38

DISTRIBUTION OF C.O.E ESTIMATED INVERSE RESMIN COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.10000	13	21	2.1000	18
2	0.20000	53	22	2.2000	19
3	0.30000	55	23	2.3000	20
4	0.40000	89	24	2.4000	11
5	0.50000	156	25	2.5000	7
6	0.60000	240	26	2.6000	20
7	0.70000	391	27	2.7000	1
8	0.80000	553	28	2.8000	10
9	0.90000	668	29	2.9000	3
10	1.00000	691	30	3.0000	5
11	1.10000	652	31	3.1000	6
12	1.20000	457	32	3.2000	5
13	1.30000	293	33	3.3000	5
14	1.40000	200	34	3.4000	1
15	1.50000	108	35	3.5000	3
16	1.60000	85	36	3.6000	4
17	1.70000	43	37	3.7000	0
18	1.80000	53	38	3.8000	2
19	1.90000	32	39	3.9000	3
20	2.00000	26	40	1000.0	40

APPENDIX C

FREQUENCY DISTRIBUTION OF COEFFICIENTS

FOR

ORIGINAL AND MODIFIED ESTIMATES

DIRECT AND INVERSE COEFFICIENTS

TABLE C-1

DISTRIBUTION OF FREQUENCIES: A72 ACTUAL DIRECT COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	518	21	0.35255988	0
2	0.017627994	4152	22	0.37018787	2
3	0.035255988	156	23	0.38781587	0
4	0.052883982	67	24	0.40544386	2
5	0.070511976	29	25	0.42307186	2
6	0.088139970	26	26	0.44069985	2
7	0.10576796	19	27	0.45832784	1
8	0.12339596	13	28	0.47595584	0
9	0.14102395	7	29	0.49358383	2
10	0.15865195	7	30	0.51121183	1
11	0.17627994	6	31	0.52883982	0
12	0.19390793	3	32	0.54646781	1
13	0.21153593	2	33	0.56409581	0
14	0.22916392	8	34	0.58172380	0
15	0.24679192	2	35	0.59935180	0
16	0.26441991	4	36	0.61697979	0
17	0.28204790	2	37	0.63460779	0
18	0.29967590	2	38	0.65223578	1
19	0.31730389	1	39	0.66986377	0
20	0.33493189	2	40	0.68749177	0

TABLE C-2

DISTRIBUTION OF FREQUENCIES NAIVE ESTIM. DIR. COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	873	21	0.45219263	3
2	0.022609632	3870	22	0.47480226	2
3	0.045219263	138	23	0.49741190	2
4	0.067828895	44	24	0.52002153	0
5	0.090438527	28	25	0.54263116	0
6	0.11304816	23	26	0.56524079	0
7	0.13565779	13	27	0.58785042	0
8	0.15826742	8	28	0.61046005	0
9	0.18087705	7	29	0.63306969	0
10	0.20348668	5	30	0.65567932	0
11	0.22609632	5	31	0.67828895	0
12	0.24870595	5	32	0.70089858	1
13	0.27131558	3	33	0.72350821	0
14	0.29392521	6	34	0.74611784	1
15	0.31653484	0	35	0.76872748	0
16	0.33914447	1	36	0.79133711	0
17	0.36175411	0	37	0.81394674	0
18	0.38436374	0	38	0.83655637	0
19	0.40697337	1	39	0.85916600	0
20	0.42958300	1	40	0.88177563	0

TABLE C-3

DISTRIBUTION OF FREQUENCIES RAS ESTIMATED DIRET COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	873	21	0.35627503	2
2	0.017813751	3801	22	0.37408878	0
3	0.035627503	160	23	0.39190253	1
4	0.053441254	63	24	0.40971628	1
5	0.071255006	30	25	0.42753004	0
6	0.089068757	18	26	0.44534379	2
7	0.10688251	19	27	0.46315754	1
8	0.12469626	10	28	0.48097129	2
9	0.14251001	13	29	0.49878504	1
10	0.16032376	5	30	0.51659879	2
11	0.17813751	6	31	0.53441254	1
12	0.19595127	6	32	0.55222630	0
13	0.21376502	3	33	0.57004005	0
14	0.23157877	5	34	0.58785380	0
15	0.24939252	5	35	0.60566755	0
16	0.26720627	1	36	0.62348130	0
17	0.28502002	3	37	0.64129505	0
18	0.30283377	3	38	0.65910880	1
19	0.32064753	1	39	0.67692256	0
20	0.33846128	1	40	0.69473631	0

TABLE C-4

DISTRIBUTION OF FREQUENCIES RECRAS ESTI. DIRECT COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	873	21	0.41956291	1
2	0.020978146	3848	22	0.44054106	1
3	0.041956291	149	23	0.46151920	3
4	0.062934437	43	24	0.48249735	1
5	0.083912582	30	25	0.50347549	2
6	0.10489073	25	26	0.52445364	1
7	0.12586887	12	27	0.54543179	0
8	0.14684702	10	28	0.56640993	0
9	0.16782516	8	29	0.58738808	0
10	0.18880331	7	30	0.60836622	0
11	0.20978146	3	31	0.62934437	0
12	0.23075960	6	32	0.65032251	0
13	0.25173775	4	33	0.67130066	1
14	0.27271589	2	34	0.69227881	0
15	0.29369404	4	35	0.71325695	1
16	0.31467218	1	36	0.73423510	0
17	0.33565033	2	37	0.75521324	0
18	0.35662848	0	38	0.77619139	0
19	0.37760662	1	39	0.79716953	0
20	0.39858477	1	40	0.81814768	0

TABLE C-5

DISTRIBUTION OF FREQUENCIES PROPVA ESTIM. DIRECT COEFFICIENT

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	873	21	0.31858601	1
2	0.015929301	3799	22	0.33451531	1
3	0.031858601	145	23	0.35044462	0
4	0.047787902	79	24	0.36637392	0
5	0.063717203	29	25	0.38230322	2
6	0.079646504	20	26	0.39823252	0
7	0.095575804	17	27	0.41416182	1
8	0.11150510	13	28	0.43009112	1
9	0.12743441	12	29	0.44602042	1
10	0.14336371	9	30	0.46194972	0
11	0.15929301	10	31	0.47787902	2
12	0.17522231	3	32	0.49380832	0
13	0.19115161	1	33	0.50973762	0
14	0.20708091	4	34	0.52566692	0
15	0.22301021	2	35	0.54159622	1
16	0.23893951	3	36	0.55752552	0
17	0.25486881	5	37	0.57345483	1
18	0.27079811	3	38	0.58938413	0
19	0.28672741	0	39	0.60531343	1
20	0.30265671	1	40	0.62124273	0

TABLE C-6

DISTRIBUTION OF FREQUENCIES FRIED ESTIM. DIRECT COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	-0.00018775698	1	21	0.35781069	2
2	0.017712165	4672	22	0.37571061	0
3	0.035612088	159	23	0.39361053	0
4	0.053512010	67	24	0.41151046	1
5	0.071411932	27	25	0.42941038	1
6	0.089311855	19	26	0.44731030	2
7	0.10721178	19	27	0.46521022	1
8	0.12511170	8	28	0.48311015	2
9	0.14301162	15	29	0.50101007	2
10	0.16091154	6	30	0.51890999	1
11	0.17881147	5	31	0.53680991	1
12	0.19671139	6	32	0.55470984	0
13	0.21461131	3	33	0.57260976	0
14	0.23251123	4	34	0.59050968	0
15	0.25041116	4	35	0.60840960	0
16	0.26831108	3	36	0.62630952	0
17	0.28621100	2	37	0.64420945	0
18	0.30411092	5	38	0.66210937	1
19	0.32201084	0	39	0.68000929	0
20	0.33991077	1	40	0.69790921	0

TABLE C-7

DISTRIBUTION OF FREQUENCIES RECLAG ESTI. DIRECT COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	873	21	0.38051969	1
2	0.019025984	3902	22	0.39954567	0
3	0.038051969	112	23	0.41857166	0
4	0.057077953	35	24	0.43759764	0
5	0.076103938	26	25	0.45662363	0
6	0.095129922	18	26	0.47564961	2
7	0.11415591	14	27	0.49467560	2
8	0.13318189	10	28	0.51370158	1
9	0.15220788	5	29	0.53272757	1
10	0.17123386	4	30	0.55175355	2
11	0.19025984	7	31	0.57077953	1
12	0.20928583	5	32	0.58980552	0
13	0.22831181	0	33	0.60883150	1
14	0.24733780	0	34	0.62785749	0
15	0.26636378	1	35	0.64688347	1
16	0.28538977	3	36	0.66590946	0
17	0.30441575	5	37	0.68493544	0
18	0.32344174	2	38	0.70396143	1
19	0.34246772	2	39	0.72298741	1
20	0.36149371	2	40	0.74201339	0

TABLE C-8

DISTRIBUTION OF FREQUENCIES RASLAG ESTI. DIRECT COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	873	21	0.35627503	2
2	0.017813751	3801	22	0.37408878	0
3	0.035627503	160	23	0.39190253	1
4	0.053441254	63	24	0.40971628	1
5	0.071255006	30	25	0.42753004	0
6	0.089068757	18	26	0.44534379	2
7	0.10688251	19	27	0.46315754	1
8	0.12469626	10	28	0.48097129	2
9	0.14251001	13	29	0.49878504	1
10	0.16032376	5	30	0.51659879	2
11	0.17813751	6	31	0.53441254	1
12	0.19595127	6	32	0.55222630	0
13	0.21376502	3	33	0.57004005	0
14	0.23157877	5	34	0.58785380	0
15	0.24939252	5	35	0.60566755	0
16	0.26720627	1	36	0.62348130	0
17	0.28502002	3	37	0.64129505	0
18	0.30283377	3	38	0.65910880	1
19	0.32064753	1	39	0.67692256	0
20	0.33846128	1	40	0.69473631	0

TABLE C-9

DISTRIBUTION OF FREQUENCIES RERALA ESTI. DIRECT COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	873	21	0.38429911	2
2	0.019214956	3830	22	0.40351407	0
3	0.038429911	151	23	0.42272902	2
4	0.057644867	52	24	0.44194398	0
5	0.076859822	30	25	0.46115893	2
6	0.096074778	22	26	0.48037389	0
7	0.11528973	16	27	0.49958885	1
8	0.13450469	9	28	0.51880380	3
9	0.15371964	10	29	0.53801876	1
10	0.17293460	6	30	0.55723371	0
11	0.19214956	9	31	0.57644867	0
12	0.21136451	0	32	0.59566362	0
13	0.23057947	5	33	0.61487858	0
14	0.24979442	2	34	0.63409353	0
15	0.26900938	5	35	0.65330849	0
16	0.28822433	2	36	0.67252345	0
17	0.30743929	3	37	0.69173840	1
18	0.32665424	0	38	0.71095336	1
19	0.34586920	2	39	0.73016831	0
20	0.36508416	0	40	0.74938327	0

TABLE C-10

DISTRIBUTION OF FREQUENCIES ALMON ESTIM. DIRECT COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	-0.28088997	1	21	0.78196039	2
2	-0.22774745	26	22	0.83510290	0
3	-0.17460494	25	23	0.88824542	0
4	-0.12146242	23	24	0.94138794	0
5	-0.068319900	49	25	0.99453046	0
6	-0.015177382	511	26	1.0476730	1
7	0.037965136	4042	27	1.1008155	1
8	0.091107654	197	28	1.1539580	0
9	0.14425017	59	29	1.2071005	0
10	0.19739269	32	30	1.2602430	0
11	0.25053521	16	31	1.3133856	2
12	0.30367773	22	32	1.3665281	0
13	0.35682024	11	33	1.4196706	0
14	0.40996276	7	34	1.4728131	0
15	0.46310528	2	35	1.5259556	0
16	0.51624780	1	36	1.5790982	1
17	0.56939032	3	37	1.6322407	0
18	0.62253283	1	38	1.6853832	-0
19	0.67567535	1	39	1.7385257	0
20	0.72881787	3	40	1.7916682	0

TABLE C-11

DISTRIBUTION OF FREQUENCIES: A72 ACTUAL INVERSE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	4.2310692E-007	1	21	1.0268588	24
2	0.051343342	4690	22	1.0782017	1
3	0.10268626	172	23	1.1295446	9
4	0.15402918	39	24	1.1808876	5
5	0.20537210	20	25	1.2322305	3
6	0.25671502	9	26	1.2835734	3
7	0.30805794	7	27	1.3349163	1
8	0.35940086	11	28	1.3862592	3
9	0.41074378	4	29	1.4376022	1
10	0.46208669	4	30	1.4889451	1
11	0.51342961	4	31	1.5402880	1
12	0.56477253	0	32	1.5916309	1
13	0.61611545	0	33	1.6429738	1
14	0.66745837	1	34	1.6943168	0
15	0.71880129	3	35	1.7456597	0
16	0.77014421	0	36	1.7970026	1
17	0.82148713	2	37	1.8483455	0
18	0.87283005	1	38	1.8996884	0
19	0.92417297	0	39	1.9510313	0
20	0.97551589	0	40	2.0023743	0

TABLE C-12

DISTRIBUTION OF FREQUENCIES NAIVE ESTI. INVERSE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	5.6816050E-007	1	21	0.92166096	1
2	0.046083588	4694	22	0.96774397	0
3	0.092166607	150	23	1.0138270	15
4	0.13824963	52	24	1.0599100	23
5	0.18433265	18	25	1.1059930	9
6	0.23041566	13	26	1.1520761	8
7	0.27649868	14	27	1.1981591	3
8	0.32258170	6	28	1.2442421	2
9	0.36866472	6	29	1.2903251	2
10	0.41474774	3	30	1.3364081	2
11	0.46083076	4	31	1.3824911	3
12	0.50691378	0	32	1.4285742	2
13	0.55299680	0	33	1.4746572	0
14	0.59907982	0	34	1.5207402	1
15	0.64516284	1	35	1.5668232	0
16	0.69124586	1	36	1.6129062	0
17	0.73732888	3	37	1.6589893	0
18	0.78341190	0	38	1.7050723	0
19	0.82949492	0	39	1.7511553	1
20	0.87557794	1	40	1.7972383	1

TABLE C-13

DISTRIBUTION OF FREQUENCIES RAS ESTIMA. INVERSE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	6.6666583E-008	1	21	1.0220450	21
2	0.051102315	4689	22	1.0731473	21
3	0.10220456	170	23	1.1242495	9
4	0.15330681	39	24	1.1753518	5
5	0.20440906	20	25	1.2264540	4
6	0.25551131	8	26	1.2775563	1
7	0.30661356	9	27	1.3286585	4
8	0.35771580	8	28	1.3797608	0
9	0.40881805	7	29	1.4308630	3
10	0.45992030	5	30	1.4819653	0
11	0.51102255	4	31	1.5330675	1
12	0.56212480	1	32	1.5841698	0
13	0.61322704	0	33	1.6352720	1
14	0.66432929	1	34	1.6863743	0
15	0.71543154	1	35	1.7374765	1
16	0.76653379	2	36	1.7885787	0
17	0.81763604	3	37	1.8396810	1
18	0.86873828	0	38	1.8907832	0
19	0.91984053	0	39	1.9418855	0
20	0.97094278	0	40	1.9929877	0

TABLE C-14

DISTRIBUTION OF FREQUENCIES RECRAS EST. INVERSE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	1.0358303E-007	1	21	0.94960658	0
2	0.047480427	4680	22	0.99708691	0
3	0.094960751	167	23	1.0445672	37
4	0.14244108	45	24	1.0920476	7
5	0.18992140	23	25	1.1395279	8
6	0.23740172	14	26	1.1870082	4
7	0.28488205	7	27	1.2344885	3
8	0.33236237	10	28	1.2819688	2
9	0.37984269	5	29	1.3294492	1
10	0.42732302	3	30	1.3769295	3
11	0.47480334	1	31	1.4244098	3
12	0.52228367	4	32	1.4718901	0
13	0.56976399	1	33	1.5193705	1
14	0.61724431	1	34	1.5668508	0
15	0.66472464	0	35	1.6143311	0
16	0.71220496	0	36	1.6618114	1
17	0.75968529	0	37	1.7092918	1
18	0.80716561	3	38	1.7567721	-0
19	0.85464593	2	39	1.8042524	1
20	0.90212626	1	40	1.8517327	0

TABLE C-15

DISTRIBUTION OF FREQUENCIES PROPVA EST. INVERSE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	2.4476103E-007	1	21	1.0795328	14
2	0.053976873	4769	22	1.1335094	8
3	0.10795350	109	23	1.1874861	9
4	0.16193013	40	24	1.2414627	1
5	0.21590676	14	25	1.2954393	3
6	0.26988339	11	26	1.3494160	2
7	0.32386001	8	27	1.4033926	1
8	0.37783664	5	28	1.4573692	1
9	0.43181327	4	29	1.5113458	0
10	0.48578990	1	30	1.5653225	0
11	0.53976653	2	31	1.6192991	0
12	0.59374316	2	32	1.6732757	0
13	0.64771978	0	33	1.7272523	0
14	0.70169641	0	34	1.7812290	0
15	0.75567304	2	35	1.8352056	1
16	0.80964967	0	36	1.8891822	0
17	0.86362630	0	37	1.9431589	1
18	0.91760293	2	38	1.9971355	0
19	0.97157955	0	39	2.0511121	0
20	1.0255562	29	40	2.1050887	0

TABLE C-16

DISTRIBUTION OF FREQUENCIES FRIED ESTI. INVERSE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	8.0060207E-008	1	21	1.0262456	27
2	0.051312355	4686	22	1.0775579	16
3	0.10262463	173	23	1.1288701	8
4	0.15393691	38	24	1.1801824	6
5	0.20524918	22	25	1.2314947	3
6	0.25656146	7	26	1.2828070	2
7	0.30787373	9	27	1.3341192	2
8	0.35918601	8	28	1.3854315	1
9	0.41049828	7	29	1.4367438	3
10	0.46181056	5	30	1.4880561	0
11	0.51312283	4	31	1.5393683	1
12	0.56443511	1	32	1.5906806	1
13	0.61574738	0	33	1.6419929	0
14	0.66705966	1	34	1.6933052	0
15	0.71837193	1	35	1.7446174	1
16	0.76968421	2	36	1.7959297	0
17	0.82099648	3	37	1.8472420	1
18	0.87230876	0	38	1.8985543	0
19	0.92362103	0	39	1.9498665	0
20	0.97493331	0	40	2.0011788	0

TABLE C-17

DISTRIBUTION OF FREQUENCIES RECLAG EST. INVERSE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	3.4823580E-008	1	21	1.0584272	36
2	0.052921393	4759	22	1.1113485	6
3	0.10584275	117	23	1.1642699	8
4	0.15876411	31	24	1.2171913	4
5	0.21168547	19	25	1.2701126	3
6	0.26460682	8	26	1.3230340	0
7	0.31752818	5	27	1.3759553	1
8	0.37044954	7	28	1.4288767	2
9	0.42337090	5	29	1.4817981	2
10	0.47629225	5	30	1.5347194	3
11	0.52921361	1	31	1.5876408	1
12	0.58213497	1	32	1.6405621	0
13	0.63505633	1	33	1.6934835	0
14	0.68797769	1	34	1.7464048	0
15	0.74089904	1	35	1.7993262	0
16	0.79382040	0	36	1.8522476	0
17	0.84674176	0	37	1.9051689	0
18	0.89966312	0	38	1.9580903	2
19	0.95258447	1	39	2.0110116	0
20	1.0055058	8	40	2.0639330	0

TABLE C-18

DISTRIBUTION OF FREQUENCIES RASLAG EST. INVERSE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	6.6666583E-008	1	21	1.0220450	21
2	0.051102315	4689	22	1.0731473	21
3	0.10220456	170	23	1.1242495	9
4	0.15330681	39	24	1.1753518	5
5	0.20440906	20	25	1.2264540	4
6	0.25551131	8	26	1.2775563	1
7	0.30661356	9	27	1.3286585	4
8	0.35771580	8	28	1.3797608	0
9	0.40881805	7	29	1.4308630	3
10	0.45992030	5	30	1.4819653	0
11	0.51102255	4	31	1.5330675	1
12	0.56212480	1	32	1.5841698	0
13	0.61322704	0	33	1.6352720	1
14	0.66432929	1	34	1.6863743	0
15	0.71543154	1	35	1.7374765	1
16	0.76653379	2	36	1.7885787	0
17	0.81763604	3	37	1.8396810	1
18	0.86873828	0	38	1.8907832	-0
19	0.91984053	0	39	1.9418855	0
20	0.97094278	0	40	1.9929877	0

TABLE C-19

DISTRIBUTION OF FREQUENCIES RERALA EST. INVERSE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	1.0363052E-007	1	21	1.0255532	27
2	0.051277759	4698	22	1.0768309	15
3	0.10255542	161	23	1.1281085	10
4	0.15383307	37	24	1.1793862	5
5	0.20511073	19	25	1.2306638	4
6	0.25638838	16	26	1.2819415	2
7	0.30766604	7	27	1.3332192	1
8	0.35894369	7	28	1.3844968	3
9	0.41022135	7	29	1.4357745	2
10	0.46149901	4	30	1.4870521	0
11	0.51277666	3	31	1.5383298	1
12	0.56405432	1	32	1.5896074	0
13	0.61533197	0	33	1.6408851	0
14	0.66660963	1	34	1.6921627	1
15	0.71788728	0	35	1.7434404	1
16	0.76916494	1	36	1.7947181	0
17	0.82044260	1	37	1.8459957	0
18	0.87172025	3	38	1.8972734	1
19	0.92299791	0	39	1.9485510	0
20	0.97427556	0	40	1.9998287	0

TABLE C-20

DISTRIBUTION OF FREQUENCIES ALMON ESTI. INVERSE COEFFICIENTS

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	-0.15896817	1	21	0.64250741	2
2	-0.11889439	17	22	0.68258119	0
3	-0.078820608	39	23	0.72265497	1
4	-0.038746829	62	24	0.76272875	1
5	0.0013269501	2226	25	0.80280253	1
6	0.041400729	2247	26	0.84287631	0
7	0.081474508	228	27	0.88295009	1
8	0.12154829	48	28	0.92302387	2
9	0.16162207	27	29	0.96309764	0
10	0.20169584	23	30	1.0031714	11
11	0.24176962	21	31	1.0432452	30
12	0.28184340	4	32	1.0833190	9
13	0.32191718	7	33	1.1233928	7
14	0.36199096	2	34	1.1634665	6
15	0.40206474	5	35	1.2035403	0
16	0.44213852	2	36	1.2436141	2
17	0.48221230	0	37	1.2836879	1
18	0.52228608	3	38	1.3237617	1
19	0.56235985	1	39	1.3638354	1
20	0.60243363	0	40	1.4039092	1

TABLE C-21

DISTRIBUTION OF FREQUENCIES:MDNAVKEY ESTIMATED DIRECT COEF.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.0000000	827	21	0.45219263	4
2	0.022609632	3903	22	0.47480226	2
3	0.045219263	149	23	0.49741190	1
4	0.067828895	44	24	0.52002153	0
5	0.090438527	29	25	0.54263116	1
6	0.11304816	24	26	0.56524079	0
7	0.13565779	14	27	0.58785042	0
8	0.15826742	8	28	0.61046005	0
9	0.18087705	5	29	0.63306969	0
10	0.20348668	6	30	0.65567932	0
11	0.22609632	5	31	0.67828895	0
12	0.24870595	4	32	0.70089858	1
13	0.27131558	2	33	0.72350821	0
14	0.29392521	6	34	0.74611784	1
15	0.31653484	0	35	0.76872748	0
16	0.33914447	2	36	0.79133711	0
17	0.36175411	0	37	0.81394674	0
18	0.38436374	0	38	0.83655637	0
19	0.40697337	1	39	0.85916600	0
20	0.42958300	1	40	0.88177563	0

TABLE C-22

DISTRIBUTION OF FREQUENCIES:MDNAVBIG ESTIMATED DIRECT COEFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.0000000	875	21	0.35255988	0
2	0.017627994	3828	22	0.37018787	2
3	0.035255988	129	23	0.38781587	0
4	0.052883982	63	24	0.40544386	2
5	0.070511976	28	25	0.42307186	2
6	0.088139970	25	26	0.44069985	2
7	0.10576796	19	27	0.45832784	1
8	0.12339596	13	28	0.47595584	0
9	0.14102395	7	29	0.49358383	2
10	0.15865195	7	30	0.51121183	1
11	0.17627994	6	31	0.52883982	0
12	0.19390793	3	32	0.54646781	1
13	0.21153593	2	33	0.56409581	0
14	0.22916392	8	34	0.58172380	0
15	0.24679192	2	35	0.59935180	0
16	0.26441991	4	36	0.61697979	0
17	0.28204790	2	37	0.63460779	0
18	0.29967590	2	38	0.65223578	1
19	0.31730389	1	39	0.66986377	0
20	0.33493189	2	40	0.68749177	0

TABLE C-23

DISTRIBUTION OF FREQUENCIES:MDNAVMIP ESTIMATED DIRECT COEFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	877	21	0.35255988	0
2	0.017627994	3817	22	0.37018787	1
3	0.035255988	138	23	0.38781587	0
4	0.052883982	65	24	0.40544386	2
5	0.070511976	27	25	0.42307186	2
6	0.088139970	30	26	0.44069985	2
7	0.10576796	15	27	0.45832784	1
8	0.12339596	14	28	0.47595584	0
9	0.14102395	5	29	0.49358383	2
10	0.15865195	8	30	0.51121183	1
11	0.17627994	6	31	0.52883982	0
12	0.19390793	2	32	0.54646781	1
13	0.21153593	3	33	0.56409581	0
14	0.22916392	7	34	0.58172380	0
15	0.24679192	2	35	0.59935180	0
16	0.26441991	3	36	0.61697979	0
17	0.28204790	4	37	0.63460779	0
18	0.29967590	2	38	0.65223578	1
19	0.31730389	1	39	0.66986377	0
20	0.33493189	1	40	0.68749177	0

TABLE C-24

DISTRIBUTION OF FREQUENCIES:MDLAGKEY ESTIMATED DIRECT COEFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	-0.00025896276	1	21	0.35703305	3
2	0.017605638	4672	22	0.37489765	0
3	0.035470238	162	23	0.39276225	0
4	0.053334839	67	24	0.41062685	1
5	0.071199440	23	25	0.42849145	1
6	0.089064040	17	26	0.44635605	2
7	0.10692864	22	27	0.46422065	1
8	0.12479324	8	28	0.48208525	1
9	0.14265784	15	29	0.49994985	2
10	0.16052244	7	30	0.51781445	1
11	0.17838704	3	31	0.53567905	1
12	0.19625164	6	32	0.55354365	1
13	0.21411624	4	33	0.57140826	0
14	0.23198084	4	34	0.58927286	0
15	0.24984545	5	35	0.60713746	0
16	0.26771005	2	36	0.62500206	0
17	0.28557465	1	37	0.64286666	0
18	0.30343925	4	38	0.66073126	1
19	0.32130385	0	39	0.67859586	0
20	0.33916845	2	40	0.69646046	0

TABLE C-25

DISTRIBUTION OF FREQUENCIES:MDLAGBIG ESTIMATED DIRECT COEFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	-0.0041059497	1	21	0.35050691	0
2	0.013624693	4625	22	0.36823755	1
3	0.031355336	182	23	0.38596819	1
4	0.049085979	79	24	0.40369883	2
5	0.066816621	30	25	0.42142948	2
6	0.084547264	26	26	0.43916012	2
7	0.10227791	19	27	0.45689076	1
8	0.12000855	16	28	0.47462141	0
9	0.13773919	8	29	0.49235205	2
10	0.15546984	8	30	0.51008269	1
11	0.17320048	6	31	0.52781333	0
12	0.19093112	3	32	0.54554398	1
13	0.20866176	2	33	0.56327462	0
14	0.22639241	7	34	0.58100526	0
15	0.24412305	2	35	0.59873590	0
16	0.26185369	3	36	0.61646655	0
17	0.27958433	4	37	0.63419719	0
18	0.29731498	2	38	0.65192783	1
19	0.31504562	1	39	0.66965848	0
20	0.33277626	2	40	0.68738912	0

TABLE C-26

DISTRIBUTION OF FREQUENCIES:MDLAGMIP ESTIMATED DIRECT COEFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	-0.00086755965	1	21	0.35212610	0
2	0.016782123	4667	22	0.36977578	1
3	0.034431806	147	23	0.38742547	1
4	0.052081489	82	24	0.40507515	2
5	0.069731172	28	25	0.42272483	2
6	0.087380855	23	26	0.44037452	2
7	0.10503054	18	27	0.45802420	1
8	0.12268022	11	28	0.47567388	0
9	0.14032990	7	29	0.49332356	2
10	0.15797959	9	30	0.51097325	1
11	0.17562927	7	31	0.52862293	0
12	0.19327895	1	32	0.54627261	1
13	0.21092864	4	33	0.56392230	0
14	0.22857832	7	34	0.58157198	0
15	0.24622800	2	35	0.59922166	0
16	0.26387769	4	36	0.61687135	0
17	0.28152737	3	37	0.63452103	0
18	0.29917705	1	38	0.65217071	1
19	0.31682673	3	39	0.66982040	0
20	0.33447642	1	40	0.68747008	0

TABLE C-27

DISTRIBUTION OF FREQUENCIES:MDRASKEY ESTIMATED DIRECT COEFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	827	21	0.35554112	3
2	0.017777056	3847	22	0.37331817	1
3	0.035554112	160	23	0.39109523	0
4	0.053331168	65	24	0.40887229	1
5	0.071108224	26	25	0.42664934	0
6	0.088885280	18	26	0.44442640	2
7	0.10666234	20	27	0.46220345	1
8	0.12443939	11	28	0.47998051	1
9	0.14221645	13	29	0.49775757	1
10	0.15999350	6	30	0.51553462	2
11	0.17777056	5	31	0.53331168	1
12	0.19554761	6	32	0.55108873	1
13	0.21332467	2	33	0.56886579	0
14	0.23110173	5	34	0.58664284	0
15	0.24887878	6	35	0.60441990	0
16	0.26665584	1	36	0.62219696	0
17	0.28443289	1	37	0.63997401	0
18	0.30220995	4	38	0.65775107	1
19	0.31998701	0	39	0.67552812	0
20	0.33776406	2	40	0.69330518	0

TABLE C-28

DISTRIBUTION OF FREQUENCIES:MDRASBIG ESTIMATED DIRECT COEFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	875	21	0.35255988	0
2	0.017627994	3808	22	0.37018787	2
3	0.035255988	149	23	0.38781587	0
4	0.052883982	63	24	0.40544386	2
5	0.070511976	28	25	0.42307186	2
6	0.088139970	25	26	0.44069985	2
7	0.10576796	19	27	0.45832784	1
8	0.12339596	13	28	0.47595584	0
9	0.14102395	7	29	0.49358383	2
10	0.15865195	7	30	0.51121183	1
11	0.17627994	6	31	0.52883982	0
12	0.19390793	3	32	0.54646781	1
13	0.21153593	2	33	0.56409581	0
14	0.22916392	8	34	0.58172380	0
15	0.24679192	2	35	0.59935180	0
16	0.26441991	4	36	0.61697979	0
17	0.28204790	2	37	0.63460779	0
18	0.29967590	2	38	0.65223578	1
19	0.31730389	1	39	0.66986377	0
20	0.33493189	2	40	0.68749177	0

TABLE C-29

DISTRIBUTION OF FREQUENCIES:MDRASMIP ESTIMATED DIRECT COEFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	878	21	0.35255988	0
2	0.017627994	3799	22	0.37018787	1
3	0.035255988	143	23	0.38781587	1
4	0.052883982	80	24	0.40544386	2
5	0.070511976	25	25	0.42307186	2
6	0.088139970	22	26	0.44069985	2
7	0.10576796	19	27	0.45832784	1
8	0.12339596	12	28	0.47595584	0
9	0.14102395	8	29	0.49358383	2
10	0.15865195	7	30	0.51121183	1
11	0.17627994	7	31	0.52883982	0
12	0.19390793	1	32	0.54646781	1
13	0.21153593	4	33	0.56409581	0
14	0.22916392	7	34	0.58172380	0
15	0.24679192	2	35	0.59935180	0
16	0.26441991	5	36	0.61697979	0
17	0.28204790	2	37	0.63460779	0
18	0.29967590	1	38	0.65223578	1
19	0.31730389	3	39	0.66986377	0
20	0.33493189	1	40	0.68749177	0

TABLE C-30

DISTRIBUTION OF FREQUENCIES:RESMIN ESTIMATED DIRECT COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	0.00000000	873	21	0.35900937	0
2	0.017950468	3809	22	0.37695983	1
3	0.035900937	154	23	0.39491030	0
4	0.053851405	62	24	0.41286077	0
5	0.071801873	26	25	0.43081124	1
6	0.089752342	25	26	0.44876171	2
7	0.10770281	17	27	0.46671218	2
8	0.12565328	8	28	0.48466264	1
9	0.14360375	12	29	0.50261311	3
10	0.16155421	6	30	0.52056358	1
11	0.17950468	5	31	0.53851405	1
12	0.19745515	6	32	0.55646452	0
13	0.21540562	3	33	0.57441499	0
14	0.23335609	3	34	0.59236545	0
15	0.25130656	6	35	0.61031592	0
16	0.26925702	2	36	0.62826639	0
17	0.28720749	2	37	0.64621686	0
18	0.30515796	5	38	0.66416733	1
19	0.32310843	1	39	0.68211780	0
20	0.34105890	2	40	0.70006826	0

TABLE C-31

DISTRIBUTION OF FREQUENCIES:MDNAVKEY ESTIMA. INVERSE COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	5.6939183E-007	1	21	0.92166419	2
2	0.046083750	4688	22	0.96774737	0
3	0.092166931	154	23	1.0138305	14
4	0.13825011	54	24	1.0599137	24
5	0.18433329	18	25	1.1059969	10
6	0.23041647	14	26	1.1520801	7
7	0.27649965	12	27	1.1981633	2
8	0.32258284	6	28	1.2442465	3
9	0.36866602	6	29	1.2903296	2
10	0.41474920	3	30	1.3364128	2
11	0.46083238	4	31	1.3824960	2
12	0.50691556	1	32	1.4285792	2
13	0.55299874	0	33	1.4746624	0
14	0.59908192	0	34	1.5207455	2
15	0.64516510	1	35	1.5668287	0
16	0.69124828	1	36	1.6129119	0
17	0.73733146	2	37	1.6589951	0
18	0.78341464	0	38	1.7050783	0
19	0.82949783	0	39	1.7511614	1
20	0.87558101	1	40	1.7972446	1

TABLE C-32

DISTRIBUTION OF FREQUENCIES:MDNAVBIG ESTIMA. INVERSE COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	4.7618299E-007	1	21	1.0255406	25
2	0.051277482	4724	22	1.0768176	19
3	0.10255449	143	23	1.1280946	7
4	0.15383149	35	24	1.1793716	5
5	0.20510850	20	25	1.2306486	3
6	0.25638551	8	26	1.2819256	3
7	0.30766251	9	27	1.3332026	1
8	0.35893952	10	28	1.3844796	3
9	0.41021653	5	29	1.4357566	1
10	0.46149353	4	30	1.4870337	1
11	0.51277054	2	31	1.5383107	1
12	0.56404754	0	32	1.5895877	1
13	0.61532455	0	33	1.6408647	1
14	0.66660156	1	34	1.6921417	0
15	0.71787856	3	35	1.7434187	0
16	0.76915557	0	36	1.7946957	1
17	0.82043257	2	37	1.8459727	0
18	0.87170958	1	38	1.8972497	0
19	0.92298659	0	39	1.9485267	0
20	0.97426359	0	40	1.9998037	0

TABLE C-33

DISTRIBUTION OF FREQUENCIES:MDNAVMIP ESTIMA. INVERSE COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	1.0876821E-007	1	21	1.0255644	28
2	0.051278325	4723	22	1.0768427	16
3	0.10255654	140	23	1.1281209	10
4	0.15383476	42	24	1.1793991	3
5	0.20511297	17	25	1.2306773	3
6	0.25639119	12	26	1.2819555	3
7	0.30766941	9	27	1.3332337	1
8	0.35894762	7	28	1.3845120	3
9	0.41022584	4	29	1.4357902	2
10	0.46150406	3	30	1.4870684	0
11	0.51278227	3	31	1.5383466	1
12	0.56406049	0	32	1.5896248	1
13	0.61533871	1	33	1.6409030	0
14	0.66661692	2	34	1.6921812	0
15	0.71789514	1	35	1.7434595	0
16	0.76917335	2	36	1.7947377	1
17	0.82045157	1	37	1.8460159	0
18	0.87172979	0	38	1.8972941	0
19	0.92300800	0	39	1.9485723	0
20	0.97428622	0	40	1.9998505	0

TABLE C-34

DISTRIBUTION OF FREQUENCIES:MDLAGKEY ESTIMA. INVERSE COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	7.4630169E-008	1	21	1.0260146	25
2	0.051300800	4697	22	1.0773153	19
3	0.10260152	163	23	1.1286160	7
4	0.15390225	39	24	1.1799168	4
5	0.20520298	19	25	1.2312175	4
6	0.25650370	9	26	1.2825182	2
7	0.30780443	8	27	1.3338189	3
8	0.35910515	6	28	1.3851197	1
9	0.41040588	9	29	1.4364204	2
10	0.46170660	5	30	1.4877211	0
11	0.51300733	4	31	1.5390218	2
12	0.56430805	1	32	1.5903226	1
13	0.61560878	0	33	1.6416233	0
14	0.66690950	1	34	1.6929240	0
15	0.71821023	1	35	1.7442247	1
16	0.76951095	1	36	1.7955255	0
17	0.82081168	3	37	1.8468262	1
18	0.87211240	1	38	1.8981269	0
19	0.92341313	0	39	1.9494276	0
20	0.97471385	0	40	2.0007284	0

TABLE C-35

DISTRIBUTION OF FREQUENCIES:MDLAGBIG ESTIMA. INVERSE COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	-0.0031375864	1	21	1.0249935	24
2	0.048268968	4675	22	1.0764001	18
3	0.099675523	183	23	1.1278066	9
4	0.15108208	38	24	1.1792132	5
5	0.20248863	25	25	1.2306197	3
6	0.25389519	9	26	1.2820263	3
7	0.30530174	7	27	1.3334328	1
8	0.35670830	11	28	1.3848394	3
9	0.40811485	4	29	1.4362459	1
10	0.45952141	3	30	1.4876525	1
11	0.51092796	5	31	1.5390591	1
12	0.56233451	0	32	1.5904656	1
13	0.61374107	0	33	1.6418722	1
14	0.66514762	1	34	1.6932787	0
15	0.71655418	3	35	1.7446853	0
16	0.76796073	0	36	1.7960918	1
17	0.81936729	2	37	1.8474984	0
18	0.87077384	1	38	1.8989049	0
19	0.92218040	0	39	1.9503115	0
20	0.97358695	0	40	2.0017180	0

TABLE C-36

DISTRIBUTION OF FREQUENCIES:MDLAGMIP ESTIMA. INVERSE COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	3.3181277E-007	1	21	1.0265400	26
2	0.051327315	4682	22	1.0778670	17
3	0.10265430	175	23	1.1291940	7
4	0.15398128	42	24	1.1805209	5
5	0.20530826	22	25	1.2318479	3
6	0.25663525	9	26	1.2831749	2
7	0.30796223	6	27	1.3345019	3
8	0.35928921	10	28	1.3858289	3
9	0.41061620	6	29	1.4371559	1
10	0.46194318	3	30	1.4884828	2
11	0.51327016	5	31	1.5398098	0
12	0.56459714	0	32	1.5911368	1
13	0.61592413	0	33	1.6424638	1
14	0.66725111	1	34	1.6937908	0
15	0.71857809	3	35	1.7451178	0
16	0.76990508	0	36	1.7964447	1
17	0.82123206	1	37	1.8477717	0
18	0.87255904	2	38	1.8990987	0
19	0.92388602	0	39	1.9504257	0
20	0.97521301	0	40	2.0017527	0

TABLE C-37

DISTRIBUTION OF FREQUENCIES:MDRASKEY ESTIMA. INVERSE COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	6.5062134E-008	1	21	1.0215417	20
2	0.051077148	4699	22	1.0726188	23
3	0.10215423	160	23	1.1236959	7
4	0.15323131	40	24	1.1747730	5
5	0.20430840	19	25	1.2258501	4
6	0.25538548	9	26	1.2769271	2
7	0.30646256	8	27	1.3280042	4
8	0.35753965	7	28	1.3790813	0
9	0.40861673	8	29	1.4301584	2
10	0.45969381	5	30	1.4812355	0
11	0.51077089	3	31	1.5323126	2
12	0.56184798	2	32	1.5833896	0
13	0.61292506	0	33	1.6344667	1
14	0.66400214	1	34	1.6855438	0
15	0.71507923	1	35	1.7366209	1
16	0.76615631	1	36	1.7876980	0
17	0.81723339	3	37	1.8387750	1
18	0.86831047	0	38	1.8898521	0
19	0.91938756	1	39	1.9409292	0
20	0.97046464	0	40	1.9920063	0

TABLE C-38

DISTRIBUTION OF FREQUENCIES:MDRASBIG ESTIMA. INVERSE COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	6.8159813E-008	1	21	1.0264776	25
2	0.051323945	4695	22	1.0778015	17
3	0.10264782	169	23	1.1291254	9
4	0.15397170	35	24	1.1804492	5
5	0.20529558	22	25	1.2317731	3
6	0.25661945	9	26	1.2830970	3
7	0.30794333	8	27	1.3344209	1
8	0.35926721	10	28	1.3857448	3
9	0.41059109	4	29	1.4370686	1
10	0.46191496	4	30	1.4883925	1
11	0.51323884	4	31	1.5397164	1
12	0.56456272	0	32	1.5910403	1
13	0.61588659	0	33	1.6423641	1
14	0.66721047	1	34	1.6936880	0
15	0.71853435	3	35	1.7450119	0
16	0.76985823	0	36	1.7963358	1
17	0.82118210	2	37	1.8476596	0
18	0.87250598	1	38	1.8989835	0
19	0.92382986	0	39	1.9503074	0
20	0.97515373	0	40	2.0016313	0

TABLE C-39

DISTRIBUTION OF FREQUENCIES:MDRASMIP ESTIMA. INVERSE COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	3.2856269E-007	1	21	1.0265184	25
2	0.051326231	4684	22	1.0778443	18
3	0.10265213	173	23	1.1291702	7
4	0.15397803	41	24	1.1804961	5
5	0.20530394	23	25	1.2318220	3
6	0.25662984	9	26	1.2831479	2
7	0.30795574	6	27	1.3344738	3
8	0.35928164	10	28	1.3857997	3
9	0.41060754	6	29	1.4371256	1
10	0.46193345	3	30	1.4884515	2
11	0.51325935	5	31	1.5397774	0
12	0.56458525	0	32	1.5911033	1
13	0.61591115	0	33	1.6424292	1
14	0.66723705	1	34	1.6937551	0
15	0.71856296	3	35	1.7450810	0
16	0.76988886	0	36	1.7964069	1
17	0.82121476	1	37	1.8477328	0
18	0.87254066	2	38	1.8990587	0
19	0.92386656	0	39	1.9503846	0
20	0.97519247	0	40	2.0017105	0

TABLE C-40

DISTRIBUTION OF FREQUENCIES:RESMIN ESTIMATED INVERSE COEFFS.

NO	CLASS	FREQUENCY	NO	CLASS	FREQUENCY
1	8.1603782E-008	1	21	1.0457038	35
2	0.052285268	4694	22	1.0979890	9
3	0.10457045	168	23	1.1502742	8
4	0.15685564	37	24	1.2025594	6
5	0.20914083	18	25	1.2548445	2
6	0.26142601	9	26	1.3071297	2
7	0.31371120	8	27	1.3594149	2
8	0.36599638	10	28	1.4117001	1
9	0.41828157	6	29	1.4639853	3
10	0.47056676	5	30	1.5162705	1
11	0.52285194	2	31	1.5685557	0
12	0.57513713	3	32	1.6208408	1
13	0.62742231	1	33	1.6731260	0
14	0.67970750	0	34	1.7254112	0
15	0.73199269	1	35	1.7776964	1
16	0.78427787	2	36	1.8299816	0
17	0.83656306	3	37	1.8822668	1
18	0.88884824	0	38	1.9345520	0
19	0.94113343	0	39	1.9868371	0
20	0.99341862	0	40	2.0391223	0

APPENDIX D

REGRESSION STATISTICS FOR

ORIGINAL AND MODIFIED ESTIMATES

DIRECT AND INVERSE COEFFICIENTS

TABLE D-1

REGRESSION STATISTICS FOR ESTIMATED NAIVE DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.241	Degrees of freedom:	5039
R-squared:	0.937	Rbar-squared:	0.936
Residual SS:	0.396	Std error of est:	0.009
F(1,5039):	74325.458	Probability of	0.000
Durbin-Watson:	1.777		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000172	0.000128	-1.348810	1.823	---	---
X1	0.982769	0.003605	272.626957	0.000	0.967733	0.96773

RESTRICTED RESIDUAL SUM OF SQUARES=.39854

TABLE D-2

REGRESSION STATISTICS FOR ESTIMATED NAIVE INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	101.104	Degrees of freedom:	5039
R-squared:	0.990	Rbar-squared:	0.990
Residual SS:	1.012	Std error of est:	0.014
F(1,5039):	498240.168	Probability of F:	0.000
Durbin-Watson:	1.537		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000916	0.000204	-4.491819	2.000	---	---
X1	0.981576	0.001391	705.861295	0.000	0.994981	0.994981

RESTRICTED RESIDUAL SUM OF SQUARES=1.05851

TABLE D-3

REGRESSION STATISTICS FOR ESTIMATED RAS DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.237	Degrees of freedom:	5039
R-squared:	0.960	Rbar-squared:	0.960
Residual SS:	0.247	Std error of est:	0.007
F(1,5039):	122251.773	Probability of F:	0.000
Durbin-Watson:	1.796		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000039	0.000101	0.386054	1.301	---	---
X1	0.994837	0.002845	349.645211	0.000	0.980007	0.980007

RESTRICTED RESIDUAL SUM OF SQUARES=0.24704

TABLE D-4

REGRESSION STATISTICS FOR ESTIMATED RAS INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	104.081	Degrees of freedom:	5039
R-squared:	0.995	Rbar-squared:	0.995
Residual SS:	0.546	Std error of est:	0.010
F(1,5039):	955236.166	Probability of F:	0.000
Durbin-Watson:	1.712		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000047	0.000150	0.315766	1.248	---	---
X1	0.998317	0.001021	977.361840	0.000	0.997373	0.997373

RESTRICTED RESIDUAL SUM OF SQUARES=0.54646

TABLE D-5

REGRESSION STATISTICS FOR ESTIMATED RECRAS DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.437	Degrees of freedom:	5039
R-squared:	0.941	Rbar-squared:	0.941
Residual SS:	0.379	Std error of est:	0.009
F(1,5039):	80642.947	Probability of F:	0.000
Durbin-Watson:	1.886		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000159	0.000125	-1.273240	1.79	---	---
X1	1.000522	0.003523	283.977	0.00	0.970149	0.970149

RESTRICTED RESIDUAL SUM OF SQUARES=0.37868

TABLE D-6

REGRESSION STATISTICS FOR ESTIMATED RECRAS INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	103.508	Degrees of freedom:	5039
R-squared:	0.988	Rbar-squared:	0.988
Residual SS:	1.196	Std error of est:	0.015
F(1,5039):	431037.862	Probability of F:	0.000
Durbin-Watson:	1.633		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000298	0.000222	-1.34194	1.82	---	---
X1	0.992403	0.001512	656.534738	0.000	0.994206	0.994206

RESTRICTED RESIDUAL SUM OF SQUARES=1.20347

TABLE D-7

REGRESSION STATISTICS FOR ESTIMATED PROPVA DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	4.981	Degrees of freedom:	5039
R-squared:	0.916	Rbar-squared:	0.916
Residual SS:	0.419	Std error of est:	0.009
F(1,5039):	54822.460	Probability of F:	0.000
Durbin-Watson:	1.922		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000167	0.000131	1.269923	1.796	---	---
X1	0.868193	0.003708	234.141965	0.000	0.956986	0.956986

RESTRICTED RESIDUAL SUM OF SQUARES=0.52788

TABLE D-8

REGRESSION STATISTICS FOR ESTIMATED PROPVA INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	97.593	Degrees of freedom:	5039
R-squared:	0.982	Rbar-squared:	0.982
Residual SS:	1.732	Std error of est:	0.019
F(1,5039):	278906.646	Probability of F:	0.000
Durbin-Watson:	1.590		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.002012	0.000267	-7.540216	2.000	---	---
X1	0.960607	0.001819	528.116130	0.000	0.991087	0.991087

RESTRICTED RESIDUAL SUM OF SQUARES=1.94489

TABLE D-9

REGRESSION STATISTICS FOR ESTIMATED FRIED DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.277	Degrees of freedom:	5039
R-squared:	0.957	Rbar-squared:	0.957
Residual SS:	0.268	Std error of est:	0.007
F(1,5039):	112775.631	Probability of F:	0.000
Durbin-Watson:	1.810		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000027	0.000105	0.258503	1.204	---	---
X1	0.996395	0.002967	335.820831	0.000	0.978381	0.978381

RESTRICTED RESIDUAL SUM OF SQUARES=0.26854

TABLE D-10

REGRESSION STATISTICS FOR ESTIMATED FRIED INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	103.768	Degrees of freedom:	5039
R-squared:	0.994	Rbar-squared:	0.994
Residual SS:	0.589	Std error of est:	0.011
F(1,5039):	882296.081	Probability of F:	0.000
Durbin-Watson:	1.739		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000096	0.000156	0.614917	1.461	---	---
X1	0.996596	0.001061	939.306170	0.000	0.997157	0.997157

RESTRICTED RESIDUAL SUM OF SQUARES=0.59048

TABLE D-11

REGRESSION STATISTICS FOR ESTIMATED RECLAG DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	7.332	Degrees of freedom:	5039
R-squared:	0.924	Rbar-squared:	0.924
Residual SS:	0.554	Std error of est:	0.010
F(1,5039):	61619.343	Probability of F:	0.000
Durbin-Watson:	1.819		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.001343	0.000151	-8.882045	2.000	---	---
X1	1.058289	0.004263	248.232438	0.000	0.961460	0.961460

RESTRICTED RESIDUAL SUM OF SQUARES=0.57896

TABLE D-12

REGRESSION STATISTICS FOR ESTIMATED RECLAG INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	113.328	Degrees of freedom:	5039
R-squared:	0.983	Rbar-squared:	0.983
Residual SS:	1.969	Std error of est:	0.020
F(1,5039):	284917.715	Probability of F:	0.000
Durbin-Watson:	1.793		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.002575	0.000285	-9.048369	2.000	---	---
X1	1.035348	0.001940	533.776840	0.000	0.991273	0.991273

RESTRICTED RESIDUAL SUM OF SQUARES=2.11069

TABLE D-13

REGRESSION STATISTICS FOR ESTIMATED RASLAG DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.237	Degrees of freedom:	5039
R-squared:	0.960	Rbar-squared:	0.960
Residual SS:	0.247	Std error of est:	0.007
F(1,5039):	122251.773	Probability of F:	0.000
Durbin-Watson:	1.796		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT:	0.000039	0.000101	0.386054	1.301	---	---
X1	0.994837	0.002845	349.645211	0.000	0.980007	0.980007

RESTRICTED RESIDUAL SUM OF SQUARES=.24704

TABLE D-14

REGRESSION STATISTICS FOR ESTIMATED RASLAG INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	104.081	Degrees of freedom:	5039
R-squared:	0.995	Rbar-squared:	0.995
Residual SS:	0.546	Std error of est:	0.010
F(1,5039):	955236.167	Probability of F:	0.000
Durbin-Watson:	1.712		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000047	0.000150	0.315766	1.248	---	---
X1	0.998317	0.001021	977.361840	0.000	0.997373	0.997373

RESTRICTED RESIDUAL SUM OF SQUARES=0.54646

TABLE D-15

REGRESSION STATISTICS FOR ESTIMATED RERALA DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.535	Degrees of freedom:	5039
R-squared:	0.945	Rbar-squared:	0.945
Residual SS:	0.360	Std error of est:	0.008
F(1,5039):	86404.139	Probability of F:	0.000
Durbin-Watson:	2.005		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000177	0.000122	-1.456130	1.855	---	---
X1	1.010109	0.003436	293.945809	0.000	0.972057	0.972057

RESTRICTED RESIDUAL SUM OF SQUARES=0.36079

TABLE D-16

REGRESSION STATISTICS FOR ESTIMATED RERALA INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	105.269	Degrees of freedom:	5039
R-squared:	0.992	Rbar-squared:	0.992
Residual SS:	0.850	Std error of est:	0.013
F(1,5039):	619317.415	Probability of F:	0.000
Durbin-Watson:	1.818		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000211	0.000187	-1.126408	1.740	---	---
X1	1.002572	0.001274	786.967226	0.000	0.995956	0.995956

RESTRICTED RESIDUAL SUM OF SQUARES=0.85038

TABLE D-17

REGRESSION STATISTICS FOR ESTIMATED ALMON DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	30.968	Degrees of freedom:	5039
R-squared:	0.104	Rbar-squared:	0.104
Residual SS:	27.735	Std error of est:	0.074
F(1,5039):	587.299	Probability of F:	0.000
Durbin-Watson:	1.555		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000430	0.001069	0.402528	1.313	---	---
X1	0.730842	0.030157	24.234243	2.000	0.323086	0.323086

RESTRICTED RESIDUAL SUM OF SQUARES=28.18645

TABLE D-18

REGRESSION STATISTICS FOR ESTIMATED ALMON INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	90.620	Degrees of freedom:	5039
R-squared:	0.887	Rbar-squared:	0.887
Residual SS:	10.264	Std error of est:	0.045
F(1,5039):	39450.427	Probability of F:	0.000
Durbin-Watson:	1.506		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.002328	0.000650	-3.584187	2.000	---	---
X1	0.879497	0.004428	198.621316	0.000	0.941667	0.941667

RESTRICTED RESIDUAL SUM OF SQUARES=11.95240

TABLE D-19

REGRESSION STATISTICS FOR ESTIMATED MDNAVKEY DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.486	Degrees of freedom:	5039
R-squared:	0.949	Rbar-squared:	0.949
Residual SS:	0.330	Std error of est:	0.008
F(1,5039):	93961.218	Probability of F:	0.000
Durbin-Watson:	1.801		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000239	0.000117	-2.048618	1.959	---	---
X1	1.008557	0.003290	306.530942	0.000	0.974218	0.974218

RESTRICTED RESIDUAL SUM OF SQUARES=0.33073

TABLE D-20

REGRESSION STATISTICS FOR ESTIMATED MDNAVKEY INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	102.084	Degrees of freedom:	5039
R-squared:	0.991	Rbar-squared:	0.991
Residual SS:	0.911	Std error of est:	0.013
F(1,5039):	559499.538	Probability of F:	0.000
Durbin-Watson:	1.524		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000570	0.000194	-2.945166	1.997	---	---
X1	0.986860	0.001319	747.997017	0.000	0.995527	0.995527

RESTRICTED RESIDUAL SUM OF SQUARES=0.93384

TABLE D-21

REGRESSION STATISTICS FOR ESTIMATED MDNAVBIG DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.035	Degrees of freedom:	5039
R-squared:	0.994	Rbar-squared:	0.994
Residual SS:	0.035	Std error of est:	0.003
F(1,5039):	865693.088	Probability of F:	0.000
Durbin-Watson:	1.730		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000316	0.000038	-8.337033	2.000	---	---
X1	0.995724	0.001070	930.426294	0.000	0.997102	0.997102

RESTRICTED RESIDUAL SUM OF SQUARES=0.03565

TABLE D-22

REGRESSION STATISTICS FOR ESTIMATED MDNAVBIG INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	103.380	Degrees of freedom:	5039
R-squared:	0.999	Rbar-squared:	0.999
Residual SS:	0.088	Std error of est:	0.004
F(1,5039):	5885707.068	Probability of F:	0.000
Durbin-Watson:	1.341		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.001114	0.000060	-18.466133	2.000	---	---
X1	0.997142	0.000411	2426.047623	0.000	0.999572	0.999572

RESTRICTED RESIDUAL SUM OF SQUARES=0.09654

TABLE D-23

REGRESSION STATISTICS FOR ESTIMATED MDNAVMIP DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variab:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	5.860	Degrees of freedom:	5039
R-squared:	0.983	Rbar-squared:	0.983
Residual SS:	0.098	Std error of est:	0.004
F(1,5039):	296594.679	Probability of F:	0.000
Durbin-Watson:	1.827		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000171	0.000064	-2.697817	1.993	---	---
X1	0.975719	0.001792	544.605067	0.000	0.991612	0.991612

RESTRICTED RESIDUAL SUM OF SQUARES=0.10209

TABLE D-24

REGRESSION STATISTICS FOR ESTIMATED MDNAVMIP INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	100.609	Degrees of freedom:	5039
R-squared:	0.997	Rbar-squared:	0.997
Residual SS:	0.268	Std error of est:	0.007
F(1,5039):	1886708.381	Probability of F:	0.000
Durbin-Watson:	1.486		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.001072	0.000105	-10.209912	2.000	---	---
X1	0.982797	0.000716	1373.575037	0.000	0.998667	0.998667

RESTRICTED RESIDUAL SUM OF SQUARES=0.31152

TABLE D-25

REGRESSION STATISTICS FOR ESTIMATED MDLAGKEY DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.405	Degrees of freedom:	5039
R-squared:	0.964	Rbar-squared:	0.964
Residual SS:	0.233	Std error of est:	0.007
F(1,5039):	133555.657	Probability of F:	0.000
Durbin-Watson:	1.823		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000075	0.000098	-0.763281	1.555	---	---
X1	1.009915	0.002763	365.452674	0.000	0.981653	0.981653

RESTRICTED RESIDUAL SUM OF SQUARES=0.23348

TABLE D-26

REGRESSION STATISTICS FOR ESTIMATED MDLAGKEY INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	104.623	Degrees of freedom:	5039
R-squared:	0.995	Rbar-squared:	0.995
Residual SS:	0.511	Std error of est:	0.010
F(1,5039):	1026357.526	Probability of F:	0.000
Durbin-Watson:	1.762		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000031	0.000145	-0.213053	1.169	---	---
X1	1.001094	0.000988	1013.093049	0.000	0.997554	0.997554

RESTRICTED RESIDUAL SUM OF SQUARES=0.51127

TABLE D-27

REGRESSION STATISTICS FOR ESTIMATED MDLAGBIG DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.038	Degrees of freedom:	5039
R-squared:	0.995	Rbar-squared:	0.995
Residual SS:	0.033	Std error of est:	0.003
F(1,5039):	912584.201	Probability of F:	0.000
Durbin-Watson:	1.910		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000030	0.000037	0.802832	1.578	---	---
X1	0.996065	0.001043	955.292730	0.000	0.997251	0.997251

RESTRICTED RESIDUAL SUM OF SQUARES=0.03325

TABLE D-28

REGRESSION STATISTICS FOR ESTIMATED MDLAGBIG INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	103.718	Degrees of freedom:	5039
R-squared:	0.999	Rbar-squared:	0.999
Residual SS:	0.064	Std error of est:	0.004
F(1,5039):	8119760.544	Probability of F:	0.000
Durbin-Watson:	1.709		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000042	0.000051	0.824919	1.591	---	---
X1	0.998888	0.000351	2849.519353	0.000	0.999690	0.999690

RESTRICTED RESIDUAL SUM OF SQUARES=0.06445

TABLE D-29

REGRESSION STATISTICS FOR ESTIMATED MDLAGMIP DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.114	Degrees of freedom:	5039
R-squared:	0.984	Rbar-squared:	0.984
Residual SS:	0.101	Std error of est:	0.004
F(1,5039):	300691.281	Probability of F:	0.000
Durbin-Watson:	1.858		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000024	0.000064	0.378499	1.295	---	---
X1	0.996766	0.001818	548.353245	0.000	0.991725	0.991725

RESTRICTED RESIDUAL SUM OF SQUARES=0.10083

TABLE D-30

REGRESSION STATISTICS FOR ESTIMATED MDLAGMIP INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	104.312	Degrees of freedom:	5039
R-squared:	0.998	Rbar-squared:	0.998
Residual SS:	0.212	Std error of est:	0.006
F(1,5039):	2477310.768	Probability of F:	0.000
Durbin-Watson:	1.680		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000077	0.000093	-0.829153	1.593	---	---
X1	1.001039	0.000636	1573.947511	0.000	0.998985	0.998985

RESTRICTED RESIDUAL SUM OF SQUARES=0.21187

TABLE D-31

REGRESSION STATISTICS FOR ESTIMATED MDRASKEY DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.352	Degrees of freedom:	5039
R-squared:	0.966	Rbar-squared:	0.966
Residual SS:	0.216	Std error of est:	0.007
F(1,5039):	142937.292	Probability of F:	0.000
Durbin-Watson:	1.818		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000052	0.000094	-0.549582	1.417	---	---
X1	1.006880	0.002663	378.070485	0.000	0.982826	0.982826

RESTRICTED RESIDUAL SUM OF SQUARES=0.21658

TABLE D-32

REGRESSION STATISTICS FOR ESTIMATED MDRASKEY INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	104.813	Degrees of freedom:	5039
R-squared:	0.995	Rbar-squared:	0.995
Residual SS:	0.474	Std error of est:	0.010
F(1,5039):	1108173.647	Probability of F:	0.000
Durbin-Watson:	1.736		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000061	0.000140	-0.436837	1.338	---	---
X1	1.002184	0.000952	1052.698270	0.000	0.997734	0.997734

RESTRICTED RESIDUAL SUM OF SQUARES=0.47494

TABLE D-33

REGRESSION STATISTICS FOR ESTIMATED MDRASBIG DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.037	Degrees of freedom:	5039
R-squared:	0.995	Rbar-squared:	0.995
Residual SS:	0.032	Std error of est:	0.003
F(1,5039):	932455.299	Probability of F:	0.000
Durbin-Watson:	1.873		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000030	0.000037	0.814427	1.585	---	---
X1	0.996051	0.001031	965.637250	0.000	0.997309	0.997309

RESTRICTED RESIDUAL SUM OF SQUARES=0.03254

TABLE D-34

REGRESSION STATISTICS FOR ESTIMATED MDRASBIG INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	103.722	Degrees of freedom:	5039
R-squared:	0.999	Rbar-squared:	0.999
Residual SS:	0.064	Std error of est:	0.004
F(1,5039):	8172637.552	Probability of F:	0.000
Durbin-Watson:	1.671		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000046	0.000051	0.892248	1.628	---	---
X1	0.998908	0.000349	2858.782530	0.000	0.999692	0.999692

RESTRICTED RESIDUAL SUM OF SQUARES=0.06404

TABLE D-35

REGRESSION STATISTICS FOR ESTIMATED MDRASMIP DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.113	Degrees of freedom:	5039
R-squared:	0.985	Rbar-squared:	0.985
Residual SS:	0.095	Std error of est:	0.004
F(1,5039):	320685.773	Probability of F:	0.000
Durbin-Watson:	1.847		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.000021	0.000062	0.339883	1.266	---	---
X1	0.997187	0.001761	566.291244	0.000	0.992235	0.992235

RESTRICTED RESIDUAL SUM OF SQUARES=0.09461

TABLE D-36

REGRESSION STATISTICS FOR ESTIMATED MDRASMIP INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	104.48	Degrees of freedom:	5039
R-squared:	0.998	Rbar-squared:	0.998
Residual SS:	0.194	Std error of est:	0.006
F(1,5039):	2706017.008	Probability of F:	0.000
Durbin-Watson:	1.635		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000098	0.000089	-1.097674	1.728	---	---
X1	1.001956	0.000609	1644.997571	0.000	0.999070	0.999070

RESTRICTED RESIDUAL SUM OF SQUARES=0.19461

TABLE D-37

REGRESSION STATISTICS FOR ESTIMATED RESMIN DIRECT COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	6.485	Degrees of freedom:	5039
R-squared:	0.958	Rbar-squared:	0.958
Residual SS:	0.271	Std error of est:	0.007
F(1,5039):	115763.496	Probability of F:	0.000
Durbin-Watson:	1.714		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000101	0.000106	-0.953139	1.659	---	---
X1	1.013344	0.002978	340.240351	0.000	0.978921	0.978921

RESTRICTED RESIDUAL SUM OF SQUARES=0.27159

TABLE D-38

REGRESSION STATISTICS FOR ESTIMATED RESMIN INVERSE COEFFICIENTS

Valid cases:	5041	Dependent variable:	Y
Missing cases:	0	Deletion method:	Listwise
Total SS:	105.634	Degrees of freedom:	5039
R-squared:	0.994	Rbar-squared:	0.994
Residual SS:	0.637	Std error of est:	0.011
F(1,5039):	830052.950	Probability of F:	0.000
Durbin-Watson:	1.763		

Variable	Estimate	Standard Error	t value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.000170	0.000162	-1.047153	1.705	---	---
X1	1.005340	0.001103	911.072418	0.000	0.996978	0.996978

RESTRICTED RESIDUAL SUM OF SQUARES=0.64037

APPENDIX E

HISTOGRAMS OF FREQUENCY DISTRIBUTION OF

COEFFICIENTS OF EQUALITY

FOR

ORIGINAL AND MODIFIED ESTIMATES

DIRECT AND INVERSE COEFFICIENTS

HISTOGRAMS OF DISTRIB. OF COEFFICIENTS OF EQUALITY

FIGURE E-1

DIRECT ACTUAL

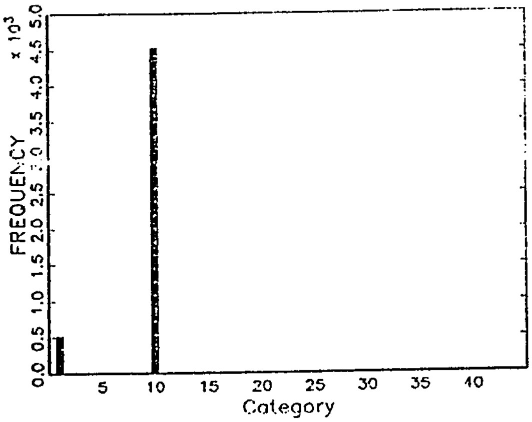


FIGURE E-2

INVERSE ACTUAL

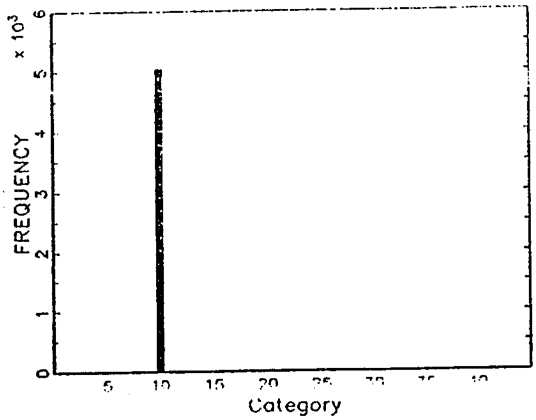


FIGURE E-3

DIRECT NAIVE

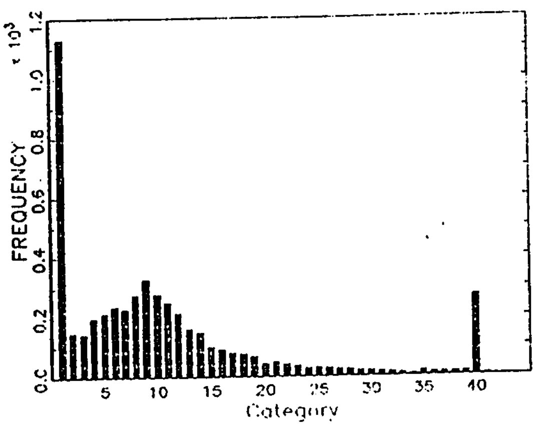
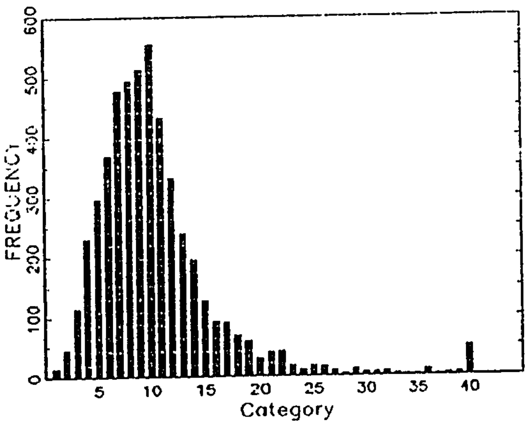


FIGURE E-4

INVERSE NAIVE



HISTOGRAMS OF DISTRIB. OF COEFFICIENTS OF EQUALITY

FIGURE E-5

DIRECT RAS

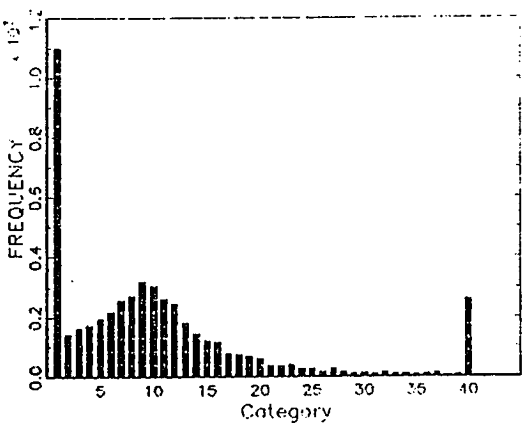


FIGURE E-6

INVERSE RAS

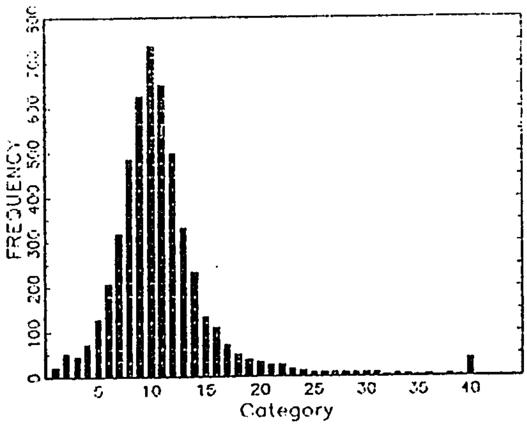


FIGURE E-7

DIRECT RECRAS

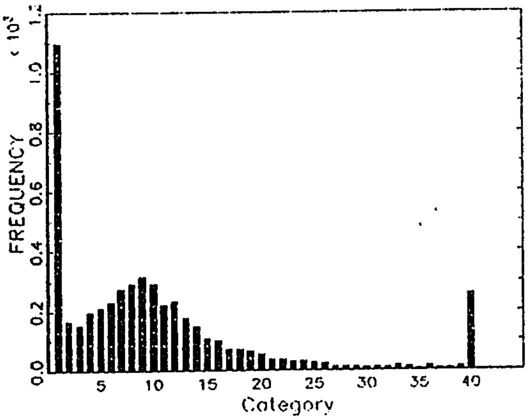
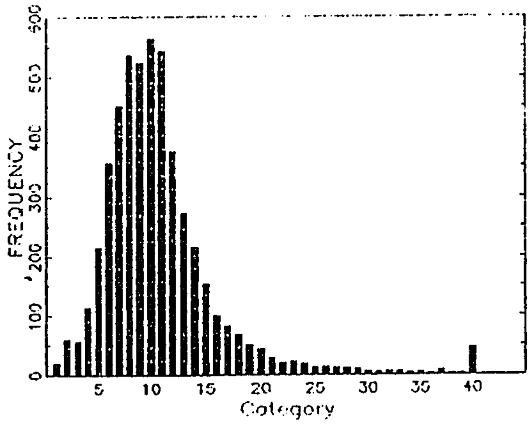


FIGURE E-8

INVERSE RECRAS



HISTOGRAMS OF DISTRIB. OF COEFFICIENTS OF EQUALITY

FIGURE E-9

DIRECT PROPVA

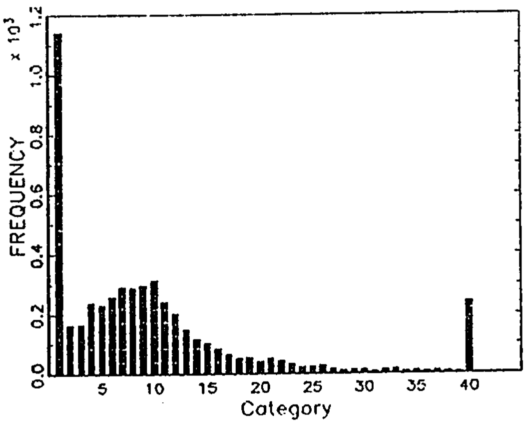


FIGURE E-10

INVERSE PROPVA

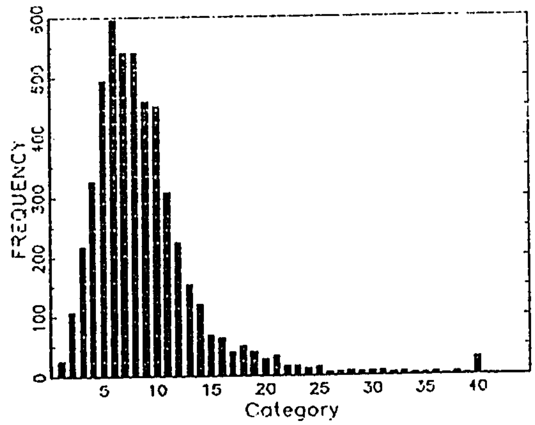


FIGURE E-11

DIRECT FRIED

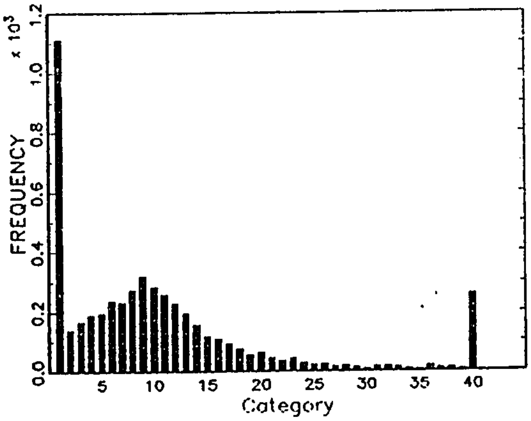
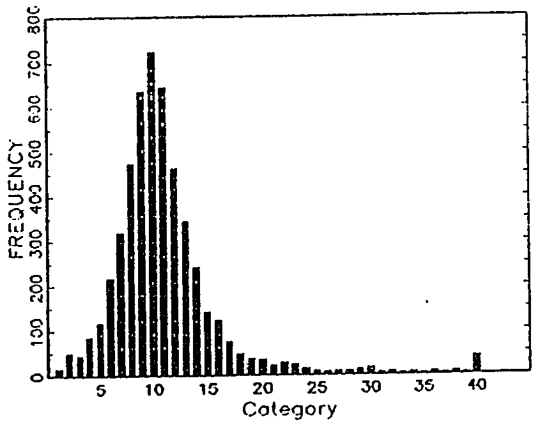


FIGURE E-12

INVERSE FRIED



HISTOGRAMS OF DISTRIB. OF COEFFICIENTS OF EQUALITY

FIGURE E-13

DIRECT RECLAG

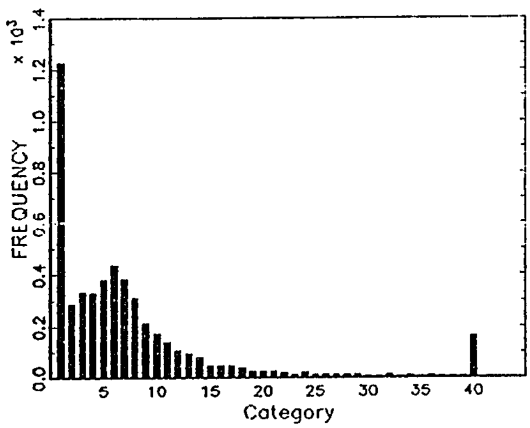


FIGURE E-14

INVERSE RECLAG

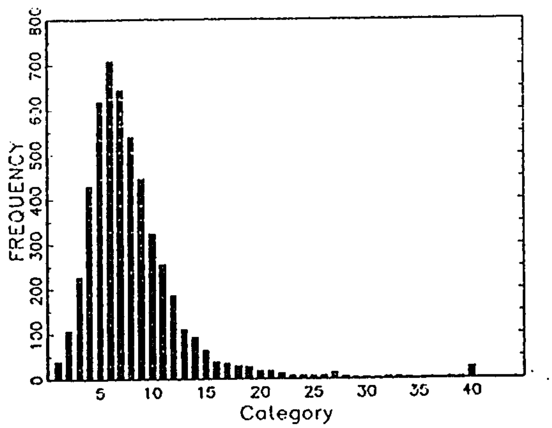


FIGURE E-15

DIRECT RASLAG

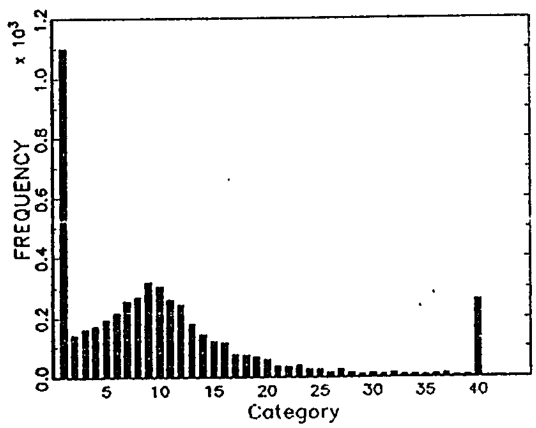
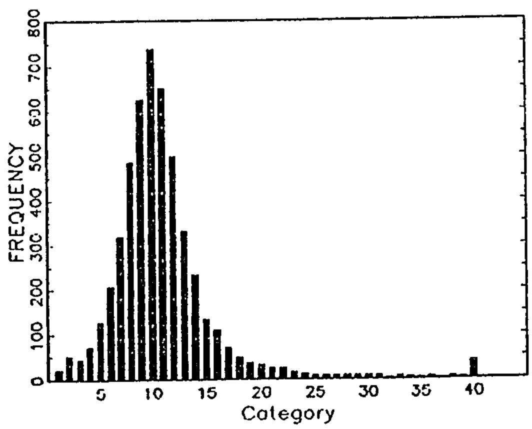


FIGURE E-16

INVERSE RASLAG



HISTOGRAMS OF DISTRIB. OF COEFFICIENTS OF EQUALITY

FIGURE E-17

DIRECT RERALA

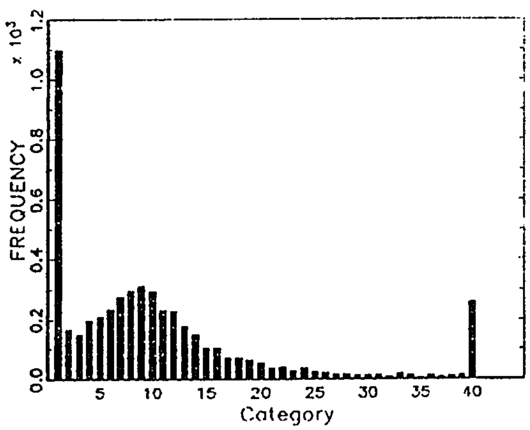


FIGURE E-18

INVERSE RERALA

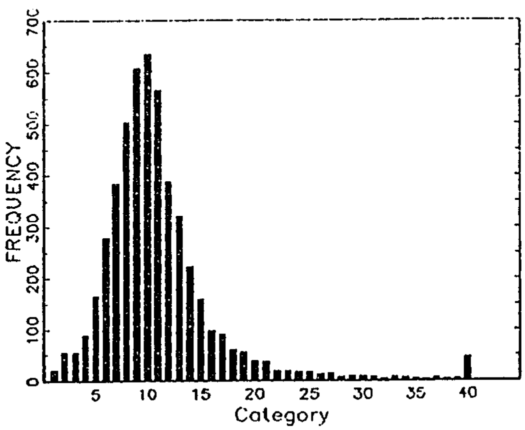


FIGURE E-19

DIRECT ALMON

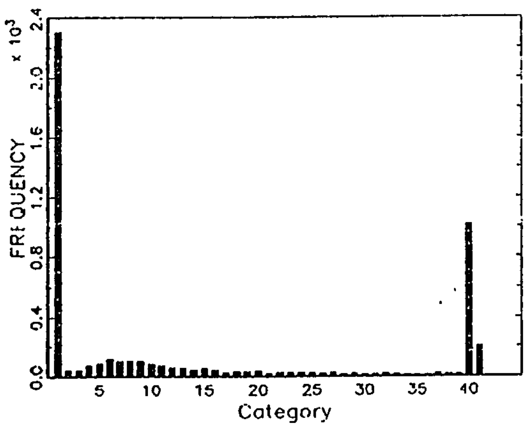
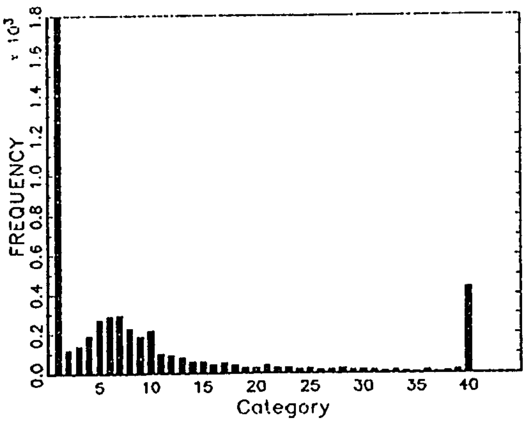


FIGURE E-20

INVERS ALMON



HISTOGRAMS OF DISTRIB. OF COEFFICIENTS OF EQUALITY

FIGURE E-21

DIRECT MDNAVKEY

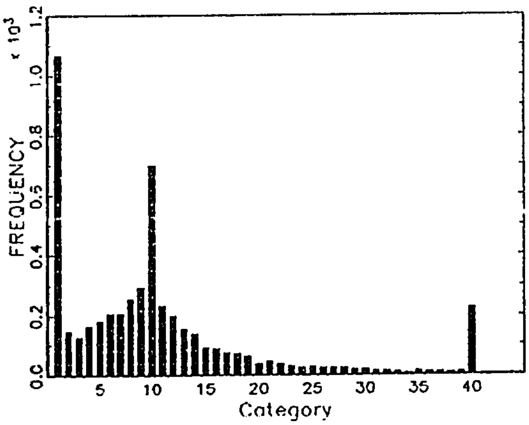


FIGURE E-22

INVERSE MDNAVKEY

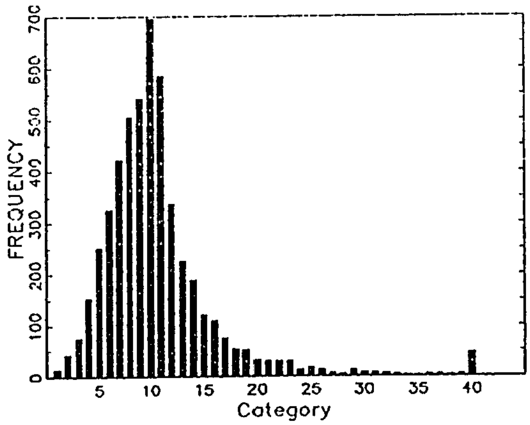


FIGURE E-23

DIRECT MDNAVBIG

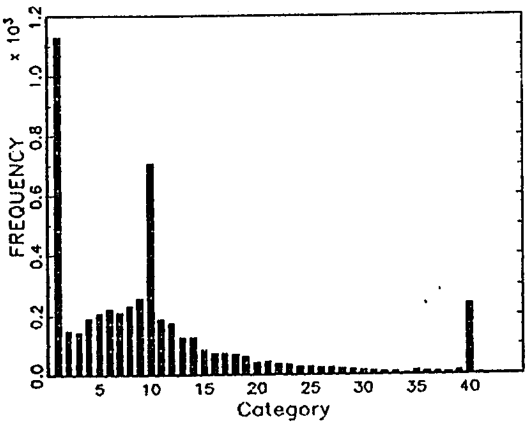
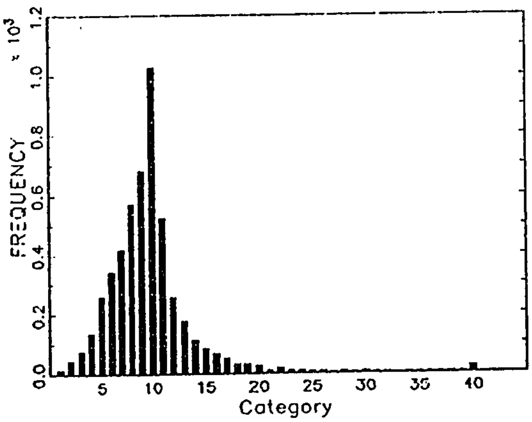


FIGURE E-24

INVERSE MDNAVBIG



HISTOGRAMS OF DISTRIB. OF COEFFICIENTS OF EQUALITY

FIGURE E-25

DIRECT MDNAVMIP

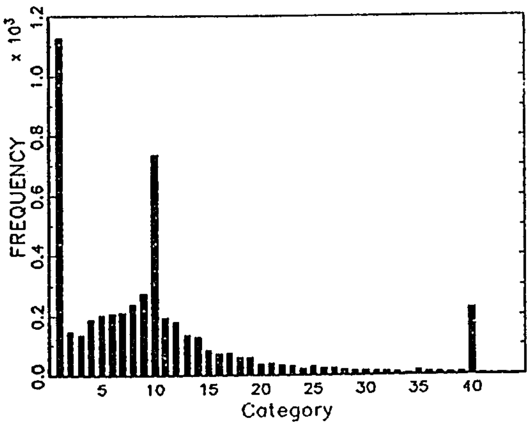


FIGURE E-26

INVERSE MDNAVMIP

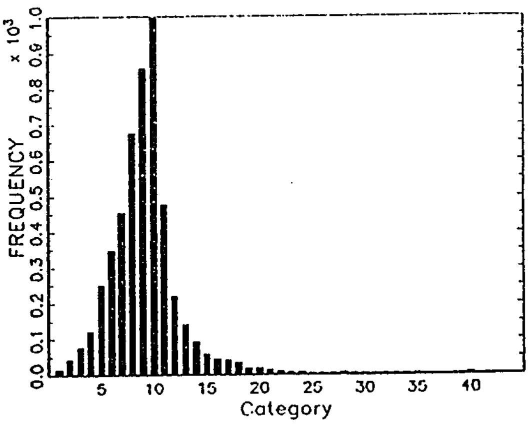


FIGURE E-27

DIRECT MDLAGKEY

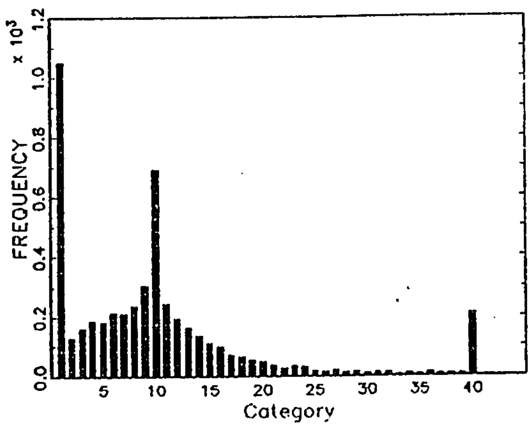
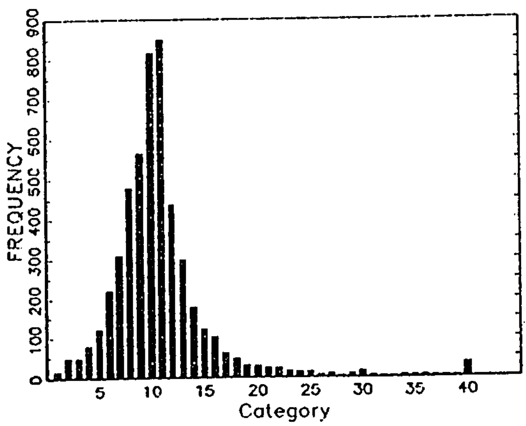


FIGURE E-28

INVERSE MDLAGKEY



HISTOGRAMS OF DISTRIB. OF COEFFICIENTS OF EQUALITY

FIGURE E-29

DIRECT MDLAGBIG

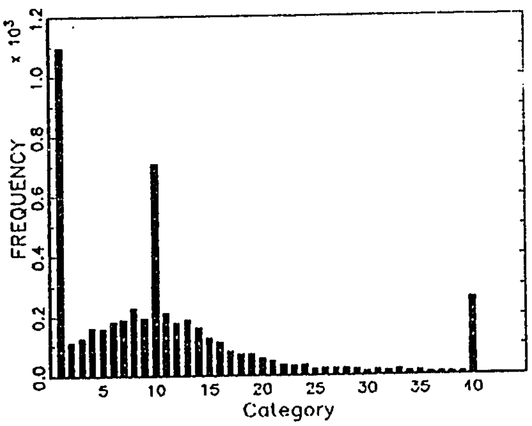


FIGURE E-30

INVERSE MDLAGBIG

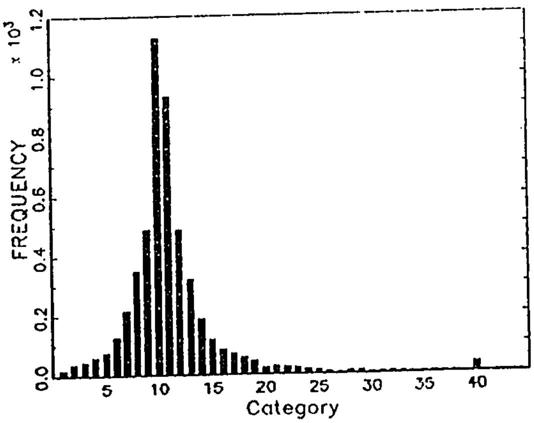


FIGURE E-31

DIRECT MDLAGMIP

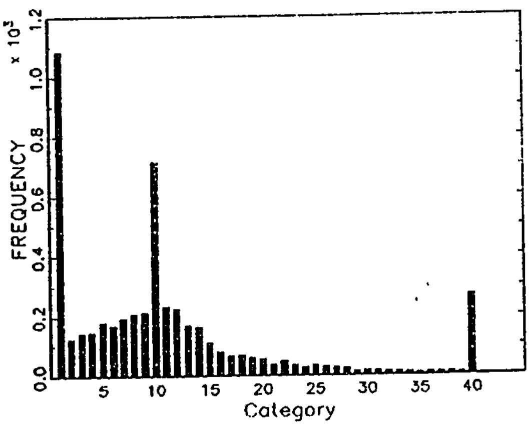
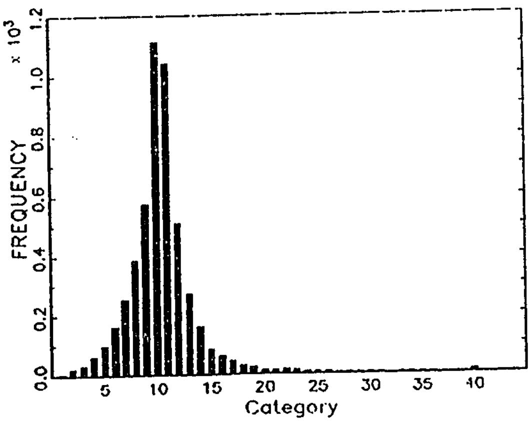


FIGURE E-32

INVERSE MDLAGMIP



HISTOGRAMS OF DISTRIB. OF COEFFICIENTS OF EQUALITY

FIGURE E-33

DIRECT MDRASKEY

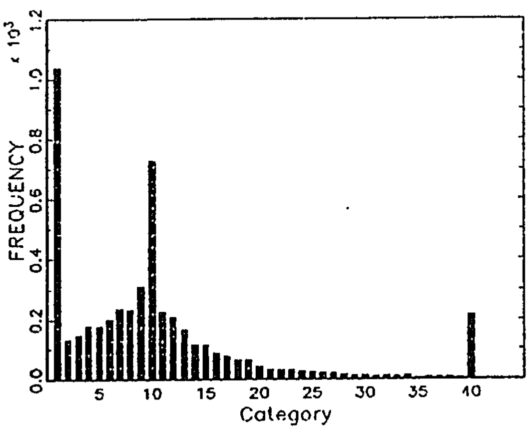


FIGURE E-34

INVERSE MDRASKEY

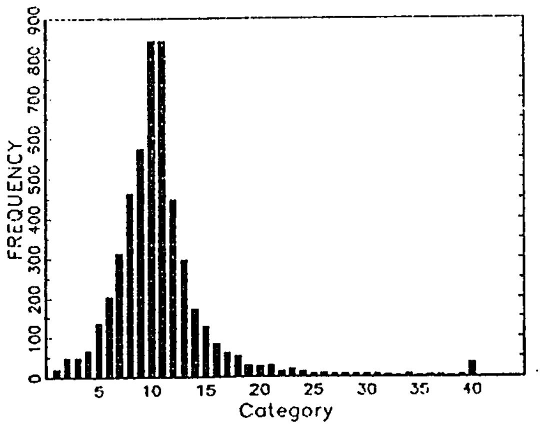


FIGURE E-35

DIRECT MDRASBIG

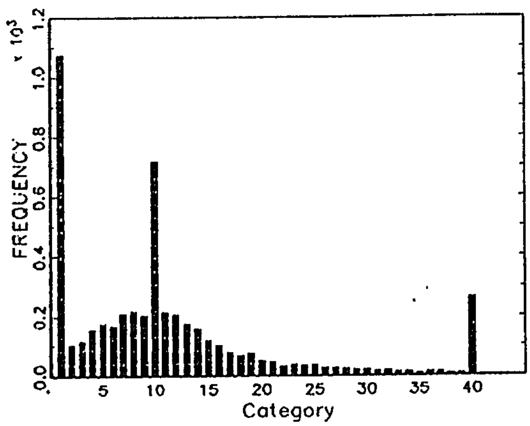
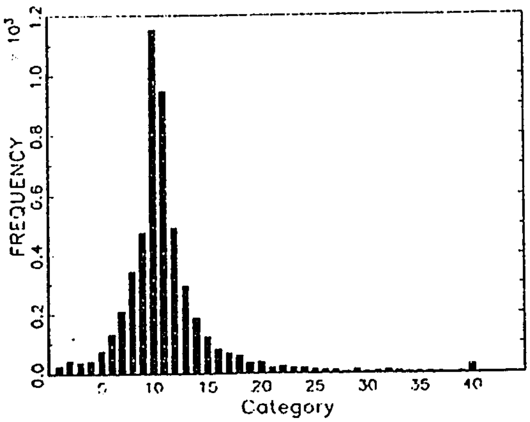


FIGURE E-36

INVERSE MDRASBIG



HISTOGRAMS OF DISTRIB. OF COEFFICIENTS OF EQUALITY

FIGURE E-37

DIRECT MDRASMIP

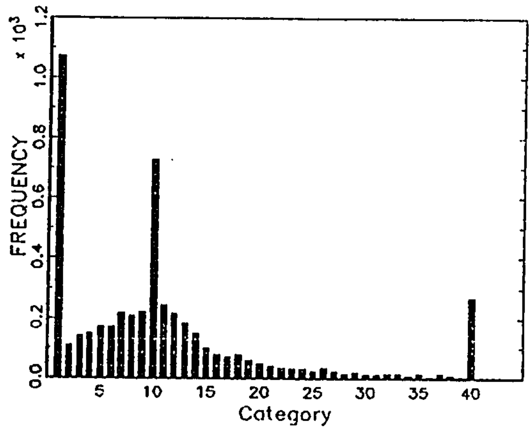


FIGURE E-38

INVERSE MDRASMIP

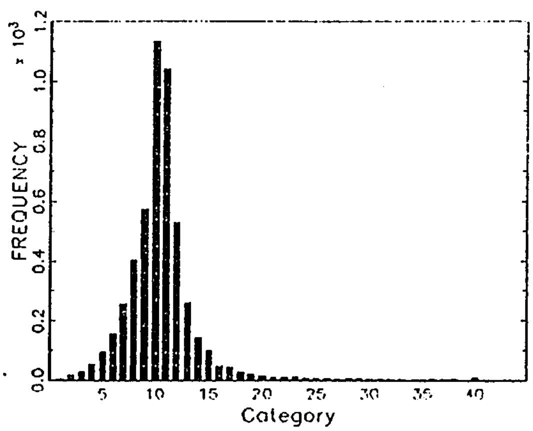


FIGURE E-39

DIRECT RESMIN

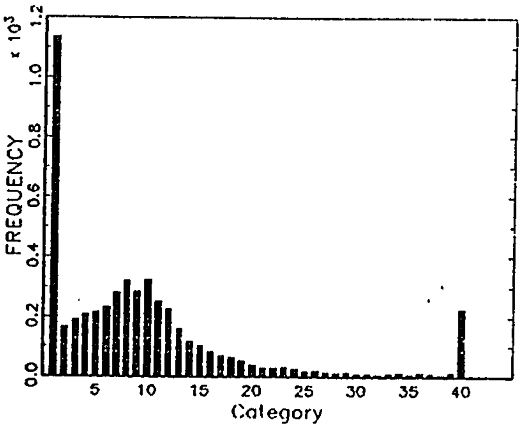
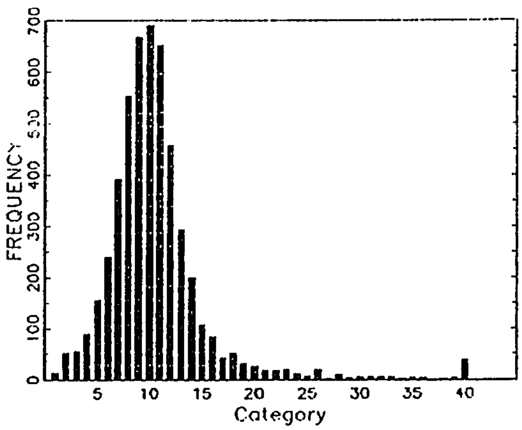


FIGURE E-40

INVERSE RESMIN



APPENDIX F

HISTOGRAMS OF FREQUENCY DISTRIBUTION OF

COEFFICIENTS

FOR

ORIGINAL AND MODIFIED ESTIMATES

DIRECT AND INVERSE COEFFICIENTS

HISTOGRAMS OF FREQUENCY DISTRIBUT. OF COEFFICIENTS

FIGURE F-1

DIRECT ACTUAL

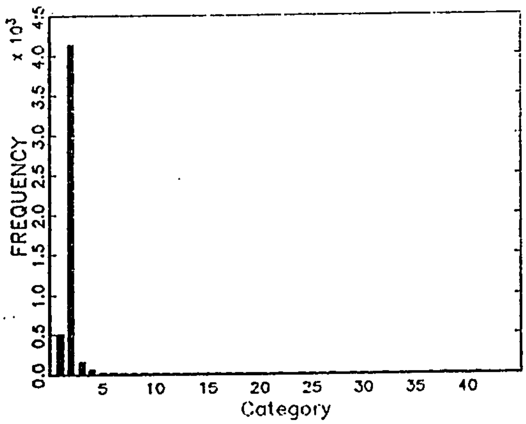


FIGURE F-2

INVERSE ACTUAL

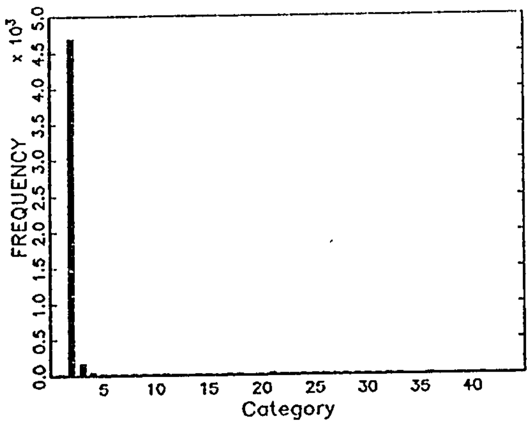


FIGURE F-3

DIRECT NAIVE

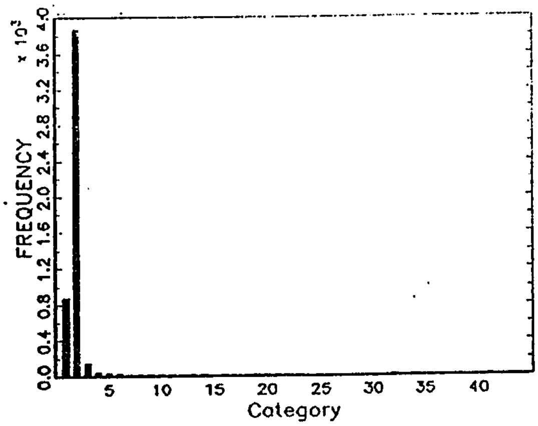
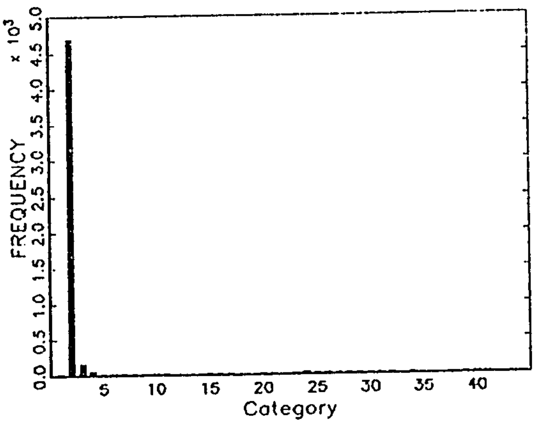


FIGURE F-4

INVERSE NAIVE



HISTOGRAMS OF FREQUENCY DISTRIBUT. OF COEFFICIENTS

FIGURE F-5

DIRECT RAS

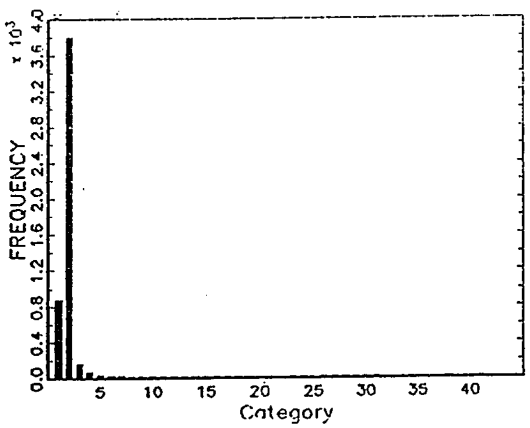


FIGURE F-6

INVERSE RAS

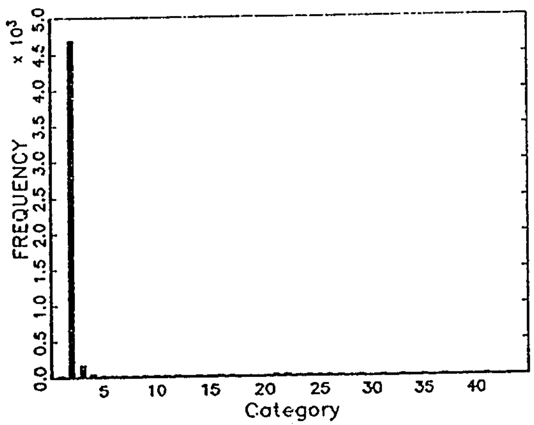


FIGURE F-7

DIRECT RECRAS

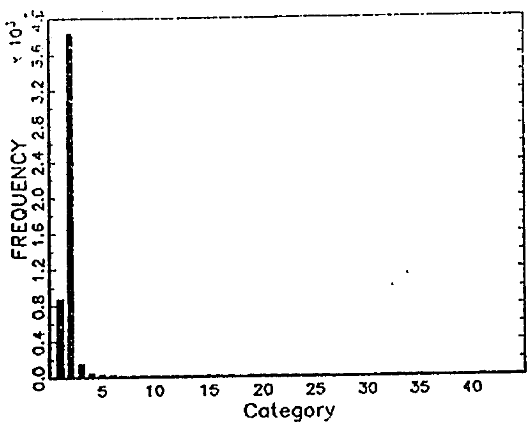
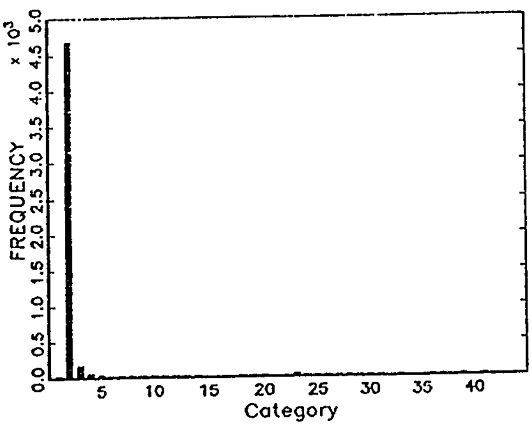


FIGURE F-8

INVERSE RECRAS



HISTOGRAMS OF FREQUENCY DISTRIBUT. OF COEFFICIENTS

FIGURE F-9

DIRECT PROPVA

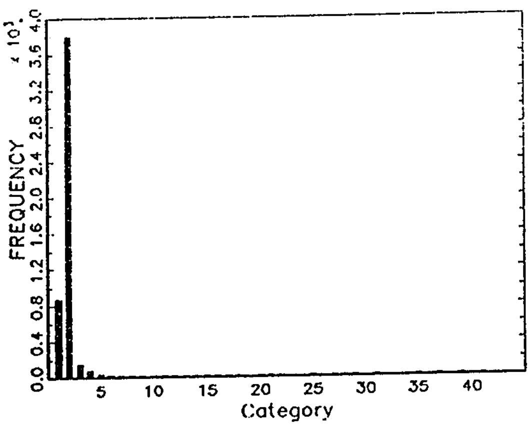


FIGURE F-10

INVERSE PROPVA

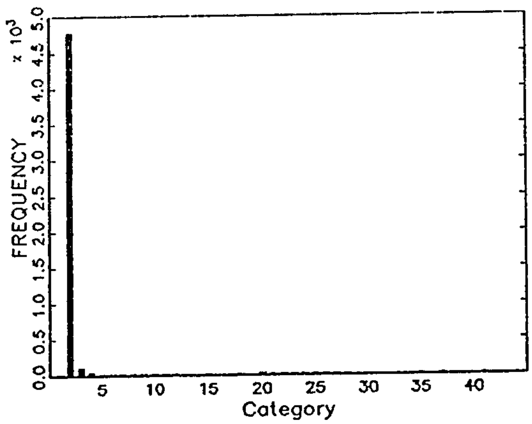


FIGURE F-11

DIRECT FRIED

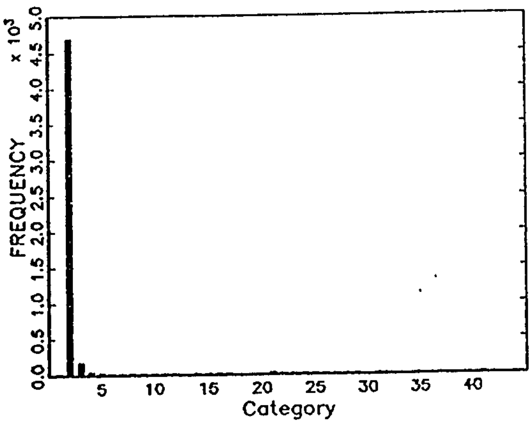
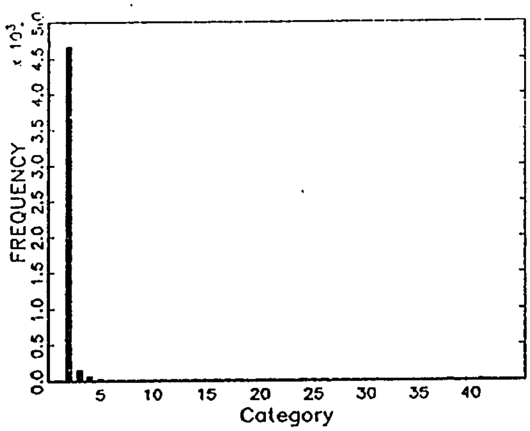


FIGURE F-12

INVERSE FRIED



HISTOGRAMS OF FREQUENCY DISTRIBUT. OF COEFFICIENTS

FIGURE F-13

DIRECT RECLAG

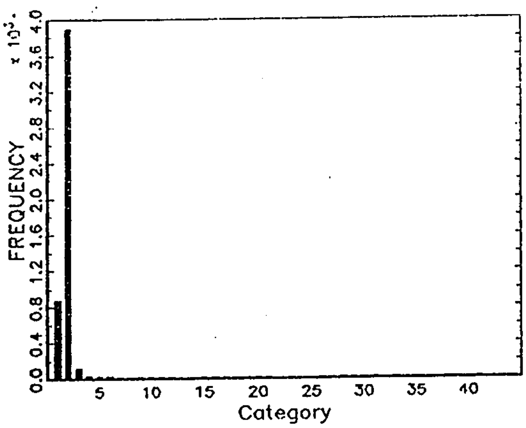


FIGURE F-14

INVERSE RECLAG

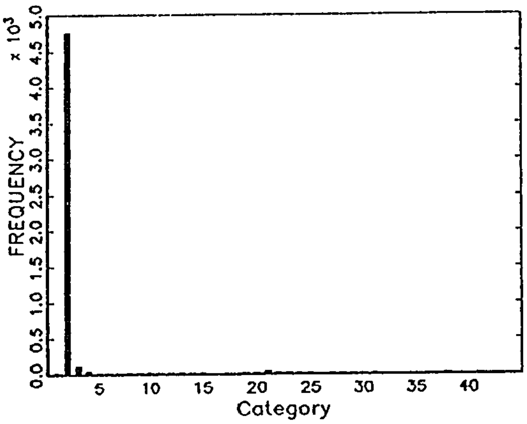


FIGURE F-15

DIRECT RASLAG

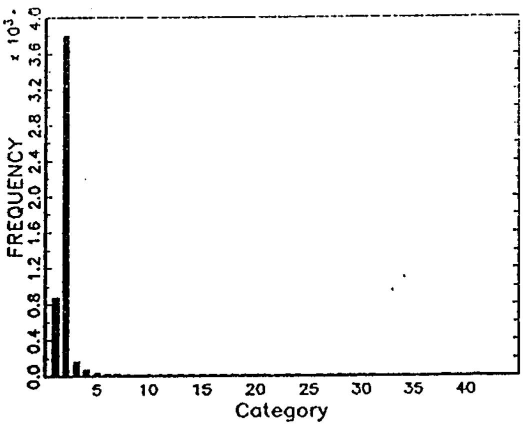
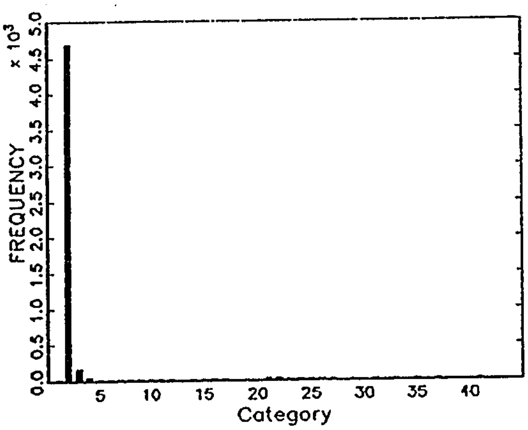


FIGURE F-16

INVERSE RASLAG



HISTOGRAMS OF FREQUENCY DISTRIBUT. OF COEFFICIENTS

FIGURE F-17

DIRECT RERALA

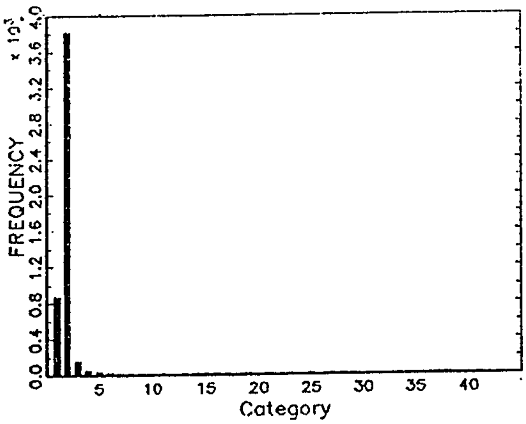


FIGURE F-18

INVERSE RERALA

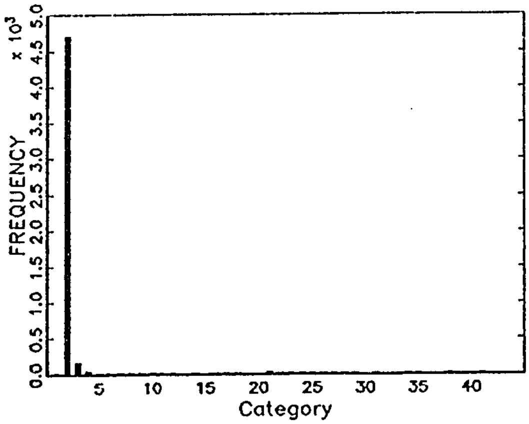


FIGURE F-19

DIRECT ALMON

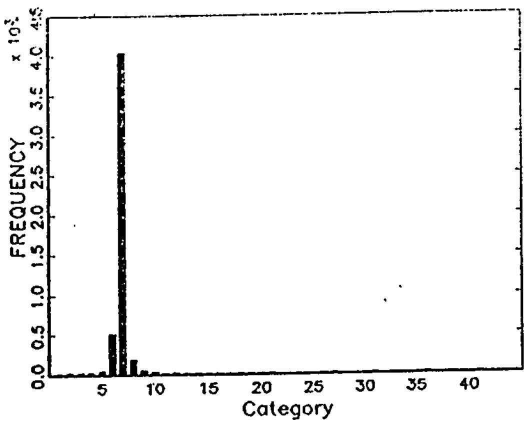
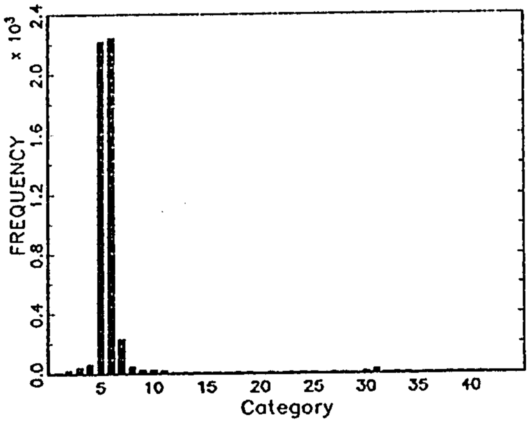


FIGURE F-20

INVERS ALMON



HISTOGRAMS OF FREQUENCY DISTRIBUT. OF COEFFICIENTS

FIGURE F-21

DIRECT MDNAVKEY

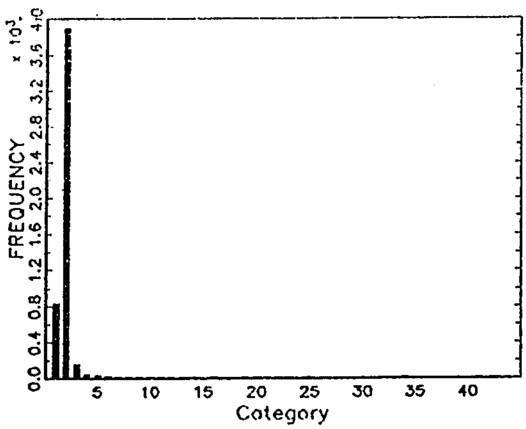


FIGURE F-22

INVERSE MDNAVKEY

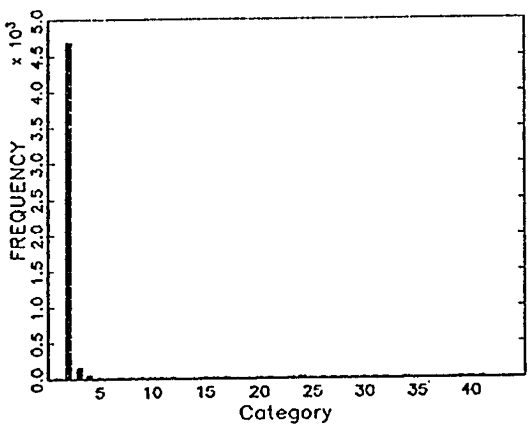


FIGURE F-23

DIRECT MDNAVBIG

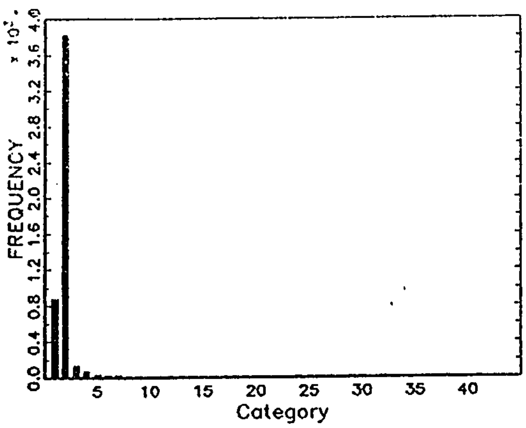
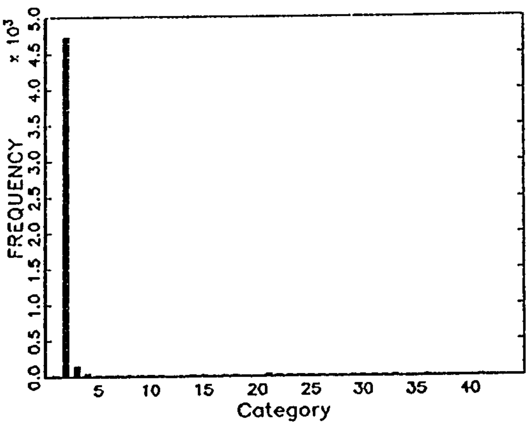


FIGURE F-24

INVERSE MDNAVBIG



HISTOGRAMS OF FREQUENCY DISTRIBUT. OF COEFFICIENTS

FIGURE F-25

DIRECT MDNAVMIP

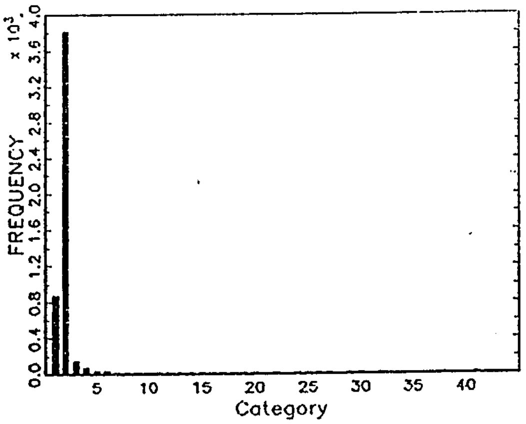


FIGURE F-26

INVERSE MDNAVMIP

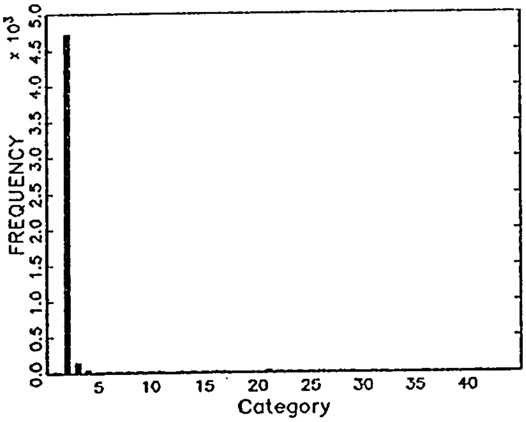


FIGURE F-27

DIRECT MDLAGKEY

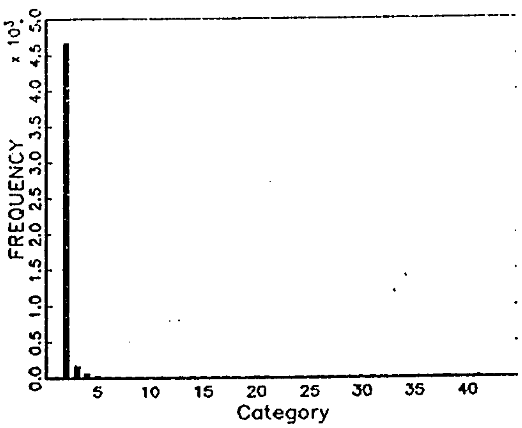
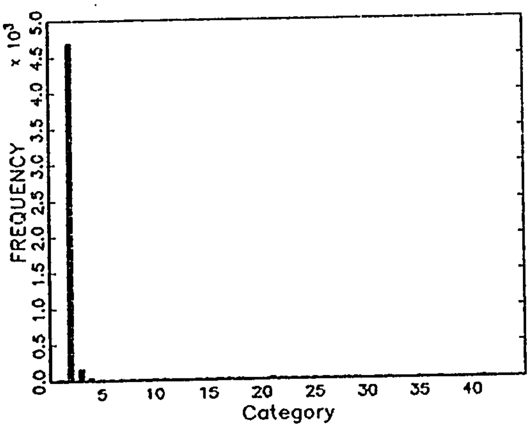


FIGURE F-28

INVERSE MDLAGKEY



HISTOGRAMS OF FREQUENCY DISTRIBUT. OF COEFFICIENTS

FIGURE F-29

DIRECT MDLAGBIG

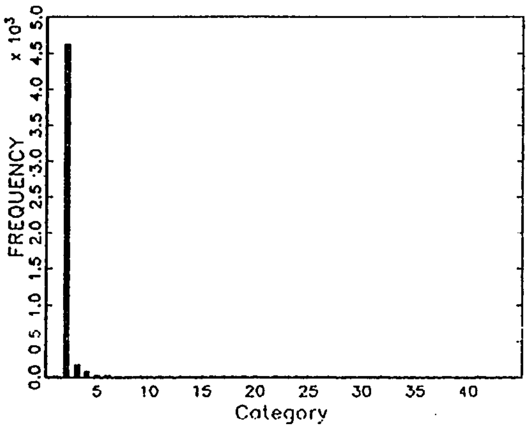


FIGURE F-30

INVERSE MDLAGBIG

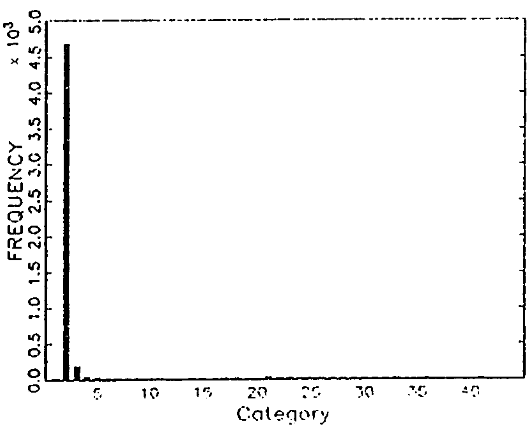


FIGURE F-31

DIRECT MDLAGMIP

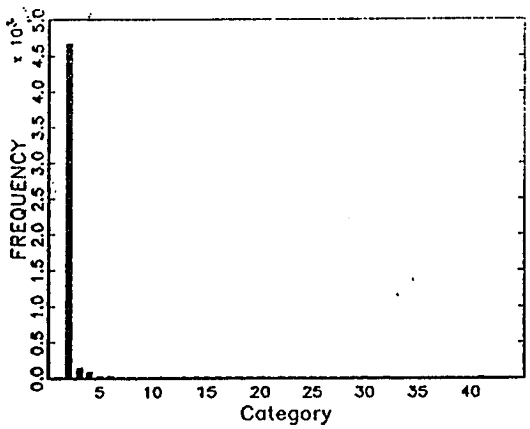
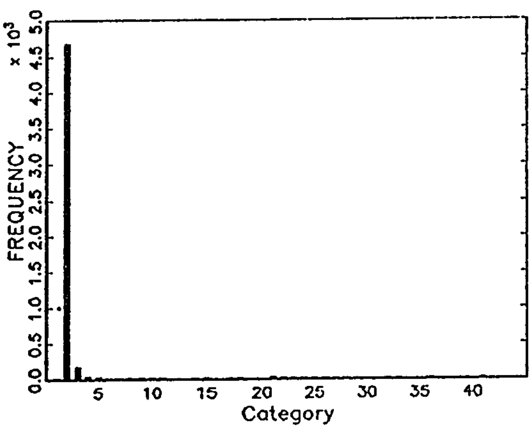


FIGURE F-32

INVERSE MDLAGMIP



HISTOGRAMS OF FREQUENCY DISTRIBUT. OF COEFFICIENTS

FIGURE F-33

DIRECT MDRASKEY

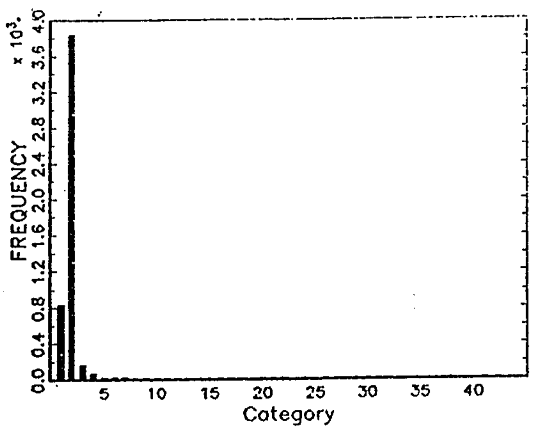


FIGURE F-34

INVERSE MDRASKEY

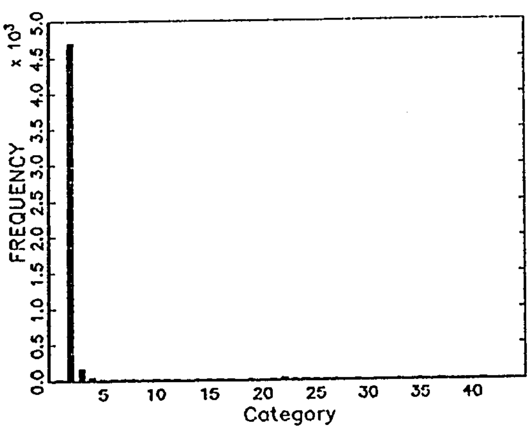


FIGURE F-35

DIRECT MDRASBIG

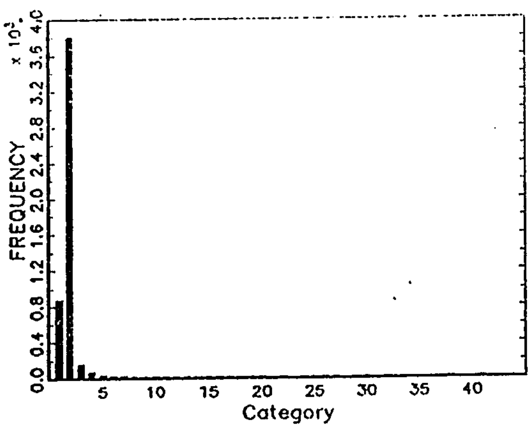
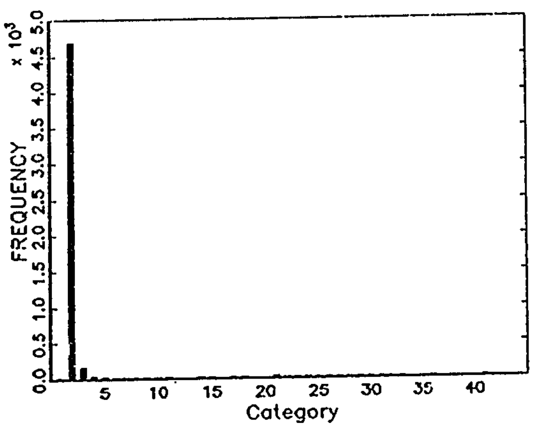


FIGURE F-36

INVERSE MDRASBIG



HISTOGRAMS OF FREQUENCY DISTRIBUT. OF COEFFICIENTS

FIGURE F-37

DIRECT MDRASMIP

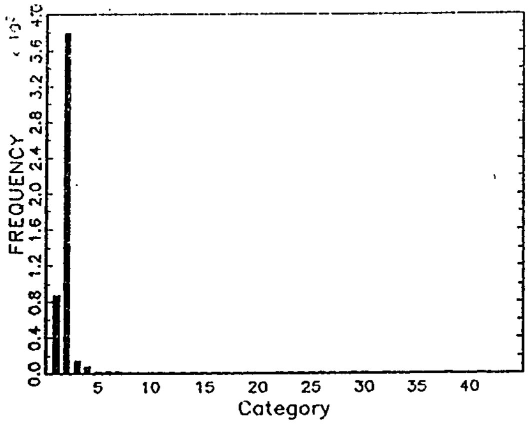


FIGURE F-38

INVERSE MDRASMIP

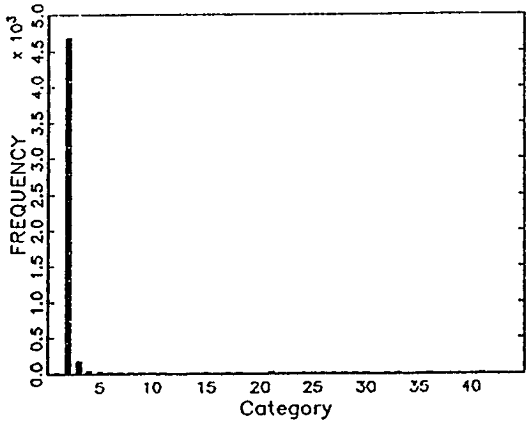


FIGURE F-39

DIRECT RESMIN

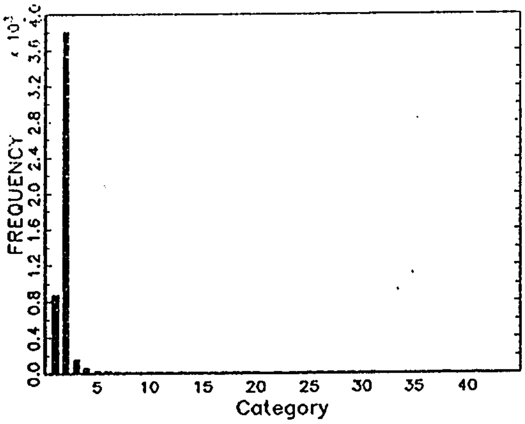
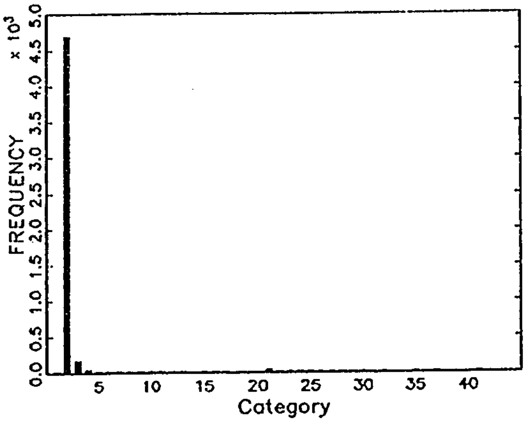


FIGURE F-40

INVERSE RESMIN



APPENDIX G

AGGREGATION SCHEMES FOR THE ORIGINAL AND

(71)(71), (35) (35), (16)(16), (6) (6)

SECTOR TABLES

TABLE G-1

SECTOR NUMBERS INCLUDED IN EACH AGGREGATION SCHEME

SECTORAL DESCRIPTION	1966 TABLE	1972 TABLE	(71)	(35)	(16)	(6)
FERROUS ORES	1	1	1	1	1	1
FERROUS METALS	2	1	1	1	1	1
NONFERROUS ORES	3	1	1	1	1	1
NONFERROUS METALS	4	1	1	1	1	1
COKE PRODUCTS	5	2	2	1	1	1
REFRACTORY-MATERIALS	6	3	3	1	1	1
INDUSTRIAL METAL PRODUCTS	7	4	4	1	1	1
COAL	8	5	5	2	2	1
OIL EXTRACTION	9	6	6	2	2	1
OIL REFINING	10	7	7	2	2	1
GAS	11	8	8	2	2	1
PEAT	12	9	9	2	2	1
OIL SHALES	13	10	10	2	2	1
ELECTRIC POWER AND STEAM	14	11	11	2	2	1
ENERGY AND POWER MACHINERY AND EQUIPMENT (M&E)	15	12	12	3	3	1

TABLE G-1 CONTINUED

SECTOR NUMBERS INCLUDED IN EACH AGGREGATION SCHEME

SECTORAL DESCRIPTION	1966 TABLE	1972 TABLE	(71)	(35)	(16)	(6)
ELECTRICAL M&E	16	13	13	3	3	1
CABLE PRODUCTS	17	14	14	3	3	1
MACHINE TOOLS	18	15	15	4	3	1
FORGING AND PRESSING M&E	19	16	16	4	3	1
CASTING M&E	20	17	17	4	3	1
TOOLS AND DIES	21	18	18	4	3	1
PRECISION INSTRUMENTS	22	19	19	5	3	1
MINING AND METALLURGICAL M&E	23	20	20	6	3	1
PUMPS AND CHEMICAL EQUIPMENT	24	21	21	7	3	1
LOGGING AND PAPER M&E	25	22	22	8	3	1
LIGHT INDUSTRY M&E	26	23	23	9	3	1
FOOD INDUSTRY M&E	27	24	24	9	3	1
PRINTING M&E	28	25	25	10	3	1

TABLE G-1 CONTINUED

SECTOR NUMBERS INCLUDED IN EACH AGGREGATION SCHEME

SECTORAL DESCRIPTION	1966 TABLE	1972 TABLE	(71)	(35)	(16)	(6)
HOISTING-TRANSPORTING M&E	29	26	26	11	3	1
CONSTRUCTION M&E	30	27	27	12	3	1
CONSTRUCTION MAT-ERIALS M&E	31	28	28	12	3	1
TRANSPORTATION M&E	32	29	29	11	3	1
AUTOMOBILES	33	30	30	11	3	1
TRACTORS AND AGRI-CULTURAL M&E	34	31	31	11	3	1
BEARINGS	35	32	32	11	3	1
RADIO AND OTHER MACHINE BUILDING	36	33	33	13	3	1
SANITARY ENGINEER-ING PRODUCTS	37	34	34	13	3	1
OTHER METAL WARES	38	35	35	13	3	1
METAL STRUCTURES	39	36	36	14	6	1
REPAIR OF M&E	40	37	37	15	3	1
ABRASIVES	41	38	38	16	4	1

TABLE G-1 CONTINUED

SECTOR NUMBERS INCLUDED IN EACH AGGREGATION SCHEME

SECTORAL DESCRIPTION	1966 TABLE	1972 TABLE	(71)	(35)	(16)	(6)
MINERAL CHEMISTRY PRODUCTS	42	39	39	16	4	1
BASIC CHEMISTRY PRODUCTS	43	40	40	16	4	1
ANILINE DYE PRODUCTS	44	41	41	16	4	1
SYNTHETIC RESINS AND PLASTICS	45	42	42	16	4	1
SYNTHETIC FIBERS	46	43	43	17	4	1
SYNTHETIC RUBBER	46	44	43	17	4	1
ORGANIC SYNTHETIC PRODUCTS	47	45	44	17	4	1
PAINTS AND LACQUERS	48	46	45	17	4	1
RUBBER AND ASBES-TOS PRODUCTS	49	47	46	18	4	1
OTHER CHEMICALS	50	48	47	16	4	1
LOGGING	51	49	48	19	5	1
SAWMILLS AND LUMBER PRODUCTS	52	50	49	19	5	1

TABLE G-1 CONTINUED

SECTOR NUMBERS INCLUDED IN EACH AGGREGATION SCHEME

SECTORAL DESCRIPTION	1966 TABLE	1972 TABLE	(71)	(35)	(16)	(6)
FURNITURE	53	51	50	20	5	1
OTHER WOODWORKING	54	52	51	20	5	1
PAPER AND PULP	55	53	52	21	5	1
WOOD CHEMISTRY PRODUCTS	56	54	53	21	5	1
CONSTRUCTION MATERIALS	57	--	54	22	6	1
CEMENT	57	55	54	22	6	1
PREFABRICATED CONCRETE	57	56	54	22	6	1
WALL MATERIALS AND TILE	57	57	54	22	6	1
ASBESTOS-CEMENT AND SLATE	57	58	54	22	6	1
ROOFING MATERIALS	57	59	54	22	6	1
CONSTRUCTION CERAMICS	57	60	54	22	6	1
OTHER CONSTRUCTION MATERIALS	57	61	54	22	6	1

TABLE G-1 CONTINUED

SECTOR NUMBERS INCLUDED IN EACH AGGREGATION SCHEME

SECTORAL DESCRIPTION	1966 TABLE	1972 TABLE	(71)	(35)	(16)	(6)
GLASS AND PORCE-LINE PRODUCTS	58	62	55	23	3	1
COTTON MATERIALS	58	63	55	23	3	1
SILK MATERIALS	58	64	55	23	3	1
WOOL MATERIALS	58	65	55	23	3	1
FLAX MATERIALS	58	66	55	23	3	1
HOSIERY AND KNITWEAR	58	67	55	23	3	1
OTHER TEXTILE PRODUCTS	58	68	55	23	3	1
TEXTILES	59	--	56	24	7	1
SEWN GOODS	60	69	57	24	7	1
OTHER LIGHT IND-USTRY PRODUCTS	61	70	58	25	9	1
FISH PRODUCTS	62	71	59	26	10	3
MEAT PRODUCTS	63	72	60	26	10	3
DAIRY PRODUCTS	64	73	61	26	10	3
SUGAR	65	74	62	27	12	3
FLOUR, BREAD, AND CONFECTIONS	66	--	63	28	12	3

TABLE G-1 CONCLUDED

SECTOR NUMBERS INCLUDED IN EACH AGGREGATION SCHEME

SECTORAL DESCRIPTION	1966 TABLE	1972 TABLE	(71)	(35)	(16)	(6)
FLOWER AND CEREALS	66	75	63	28	12	3
BREAD AND BAKERY PRODUCTS	66	76	63	28	12	3
CONFECTIONS	66	77	63	28	12	3
VEGETABLE OILS	66	78	63	28	12	3
FRUIT AND VEGETABLE PRODUCTS	66	79	63	28	12	3
OTHER FOODS	67	80	64	28	12	3
INDUSTRY NOT ELSE-WHERE CLASSIFIED	68	81	65	29	8	2
CONSTRUCTION	69	82	65	29	8	2
CROPS	70	83	66	30	13	3
ANIMAL HUSBANDRY	71	84	67	31	13	3
FORESTRY	72	85	68	32	13	3
TRANSPORTATION AND COMMUNICATIONS	73	86	69	33	14	4
TRADE AND DISTRIBUTION	74	87	70	34	15	5
OTHER BRANCHES OF MAT-ERIAL PRODUCTION	75	88	71	35	16	6

BIBLIOGRAPHY

Afrasiabi, Ahmad, and Casler, Stephen D. "Product Mix and Technological Change Within the Leontief Inverse." *Journal of Regional Science*, May 1991, Volume 31, Number 2, p 147.

Aidenoff, Abraham. "Input-Output Data in the United Nations System of National Accounts" in Leontief, W., *Contribution to Input-Output Analysis VI*, edited by Carter, A. P., and Brady, A., Amsterdam and London: North-Holland Publishing Company, 1970, pp. 349-368.

Allen, R. I. G. "Some Experiments with the RAS Method of Updating Input-Output Coefficients," *Oxford Bulletin Economics and Statistics*, volume 36, number 3, August 1974, pp. 215-228.

Allen, R.I.G., and Lecomber, J.R.C. "Some Tests of Generalized Version of RAS." Chapter 3 of Allen, R.I.G., and Gossling, W. F., (eds.) *Estimating and Projecting Input-Output Coefficients*, London: Input-Output Publishing Company, 1975.

Almon, Clopper. Jr. *Interindustry Forecasts of the American Economy*, Lexington, Mass., Lexington Books, 1974.

Almon,Clopper. Jr. "*Recent Methodological Advances in Input-Output in the United States and Canada*," Paper presented to the Forth International Conference on Input-Output Techniques, Geneva, January 1968.

Almon, Clopper. Jr. "Numerical Solution of Modified Leontief Dynamic System for Consistent Forecasting or Indicative Planning," *Econometrica*, volume 31, October 1963, pp. 665-675.

Allexeev, Michael. "*The Modified RAS Algorithm*" A research Paper for Center for International Research, Bureau of the Census, contract number BC83YAA21734, August 14, 1983.

Ara, K. "The Aggregation Problem in Input-Output Analysis," *Econometrica*, volume 27, April 1959, pp. 257-262.

Bacharach, Michael. *Biproportional Matrices and Input-Output Change*, England: Cambridge University Press, 1970.

Bacharach, Michael. "Estimating Non-Negative Matrices from Marginal Data," *International Economic Review*, volume 6, September 1965, pp. 294-310.

Bank Of Korea. Predictability Test-for Input-Output Tables, *Economic Review*, Seoul, December 1975.

Barker. T.S. "Some Experiments in Projecting Intermediate Demand" Chapter 2 of Allen, R.I.J. and Gossling, W.F., (eds.) *Estimating and Projecting Input-Output Coefficients*, London: Input-Output Publishing Company, 1975 A.

Barker, T.S. "An Analysis of the Updated 1963 Input-Output Transaction Table," chapter 4 of Allen, R.I.J. and Gossling, W.F. (eds.) *Estimating and Projecting Input-Output Coefficients*, London: Input-Output Publishing Company, 1975 B.

Barnett, H. "*Specific Industry Output Projections*," Conference on Research in Income and Wealth, Long Range Economic Projections. Princeton University Press, 1954.

Baster, J. "Stability of Trade Patterns in Regional Input-Output Tables." *Urban Studies*. Volume 17, number 1, February 1980, PP 71-75.

Belykh, A. A. "A Note on the Origin of Input-Output Analysis and the Contribution of the Early Soviet Economists: Chayanov, Bogdanov, and Kristman," *Soviet Studies*, Volume 41, Number 3. July 1989, p. 426.

Berenson, Mark L., and Levine, David M. *Basic Business Statistics*, Prentice Hall, New Jersy, 1992.

Berman, E. "*A Program for the Examination of Coefficients Variation in the Interindustry Context*," Inter-industry item Number 47, U.S. Bureau of Mines, Washington, D.C., 1953.

Beyers, William B. "On Stability of Regional Interindustry Models: the Washington Data for 1963 and 1977." *Journal of Regional Science*, Volume 12, number 3, December 1972, pp 363-374.

Bezdek, R.H. "Tests of Three Hypotheses Relating to the Leontief Input-Output Model," *Journal of Royal Statistical Society*, A, volume 147, part 3, 1984, pp.499-509.

Bezdek, R. H. "Assessing the Accuracy of Interindustry Econometric Simulations," *Economics of Planning*, volume 15, number 1, 1979, pp.

Bezdek, Roger H., and Dunham Constance R. "On the Relationship between Changes in Input-Output Coefficients and Changes in Product Mix," *Review of Economics and Statistics*, 1976, volume 58, number 3, pp. 357-379.

Bezdek, Roger H., and Dunham Constance R. "Changes in Industry Product Mix as a cause of Intertemporal Instability in Input-Output Coefficients," *Quarterly Review of Economics and Business*, 1976, Volume 16, number 1, pp. 61-76.

Bezdek, R. H., and Wending, R. M. "Current and constant dollar input-output forecasts for the US economy," *Journal of American Statistical Association*, volume 71, 1976, pp. 543-551.

Birger, E. S., Urinson, I. M., and Charnyi, V. I. "Experimental Construction of a Dynamic Input-Output Model," *Matekon*, volume 15, number 3, Spring 1979, pp. 59-82.

Bourque, Philip J. *Forecasting with Input-Output*, Graduate School of Business Administration, University of Washington, 1971.

Bullard, C. W., and Sebald, A. V. "Effects of Parametric Uncertainty and Technological Change on Input-Output Models," *Review of Economics and Statistics*, volume 59, number 1, February 1977.

Bulmer-Thomas, V. *Input-Output Analysis in Developing Countries*, John Wiley & Sons Ltd., New York, 1982.

Butterfield, Martin, and Mules, Trevor. "A Testing Routine for Evaluating Cell by Cell Accuracy in Short-cut Regional Input-Output Tables," *Journal of Regional Science*, 20, August, 1980, pp. 293-308.

Carter, A. P. *Structural Changes in the American Economy*. Harvard University Press, Cambridge, MA, 1970.

Carter, A. P., and Brody, A. (eds.) *Contributions to Input-Output Analysis*, Proceedings of the 5th International Conference on Input-Output Techniques, Amsterdam and London: North-Holland Publishing Company, 1970 (1972).

Carter, A. P., and Brody, A. (eds.) *Applications of Input-Output Analysis*, Volume 1, Proceedings of the 4th International Conference on I-O Techniques, Geneva, January 1968. North Holland Publishing Company, Amsterdam, 1970.

Carter, A. P., and Brody, A. (eds.) *Applications of Input-Output Analysis*, volume 2, Proceedings of the 4th International Conference on I-O Techniques, Geneva, January 1968. North-Holland Publishing Company, Amsterdam, 1970.

Casler, Stephen D., Afrasiabi, Ahmad, and McCauley, Michael. "Decomposing Changes in Energy Input-Output Coefficients, *Resources and Energy*, Volume 13, Number 1, April 1991, pp. 95-109.

Chenery, H. B. and Clark, P. G. *Interindustry Economics*, John Wiley, New York, 1959.

Christ,C.F. "A Review of Input-Output Analysis," in *Input-Output Analysis: An Appraisal, Studies in Income and Wealth*, National Bureau of Economic Research, volume 18, Princeton University Press, 1955, pp.137-169.

Conway, Richard S. Jr. "Changes in Regional Input-Output Coefficients and Regional Forecasting." *Regional Science and Urban Economics*. Volume 1, March 1980, pp. 153-171.

Conway, Richard S. Jr. "The Stability of Regional Input-Output Multipliers." *Environment and Planning*, number 2, February 1977, pp. 197-214.

Conway, Richard S. Jr. "A Note on the Stability of Regional Interindustry Models," *Journal of Regional Science*, volume 15, number 1 (April 1975), pp. 67-72.

Cray, Randy F. *Estimating the Effects of Structural and Price Changes in Regional Input-Output Framework*, Ph.D. dissertation, Kansas State University, 1986.

Crown William H. "An Interregional Perspective on Aggregation Bias and Information Loss in Input-Output Analysis." *Growth and Change*, 1990, volume 21, number 1(Winter):11.

Crown, William H. "An Approach to Estimating a Consistent Aggregate Input-Output Model," *Growth and Change*, volume 18, number 4, 1987, pp. 1-9

Czamanski, S., and Malizia, E. "Applicability and Limitations in the Use of National Input-Output Tables for Regional Studies." *Papers of the Regional Science Association*, 23 (1969), pp. 400-412.

Csepinszky, A., (ed.) *Input-Output Techniques: Proceedings of the Second Hungarian Conference on Input-Output Techniques*, translated by Hajdu, Gy, and Radz, J., Budapest: The Hungarian Academy of Sciences Akademia Kaido, 1976.

Davis, H. C., Lofting, E. M., and Sathaye, J. A. "A Comparison of Alternative Method of Updating Input-Output Coefficients", *Technological Forecasting and Social Change*, volume 10, number 1, 1977.

Deming, W.E., and Stephen, F.F. "On a Least Square Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known," *Annals of Mathematical Statistics*, volume 11, 1940, pp. 427-444.

Denisenko, Ivan M. "Intersectoral Balances of Production and Distribution for a National Economy: Key Aspects of the Practice in the USSR" IN: International Comparisons of interindustry Data. *Industrial Planning and Programming Series* number 2 pp.119-131

Dipasquale, Denise, and Polenske, Karen R. "Output, Income, and Employment Input-Output Multipliers" In Economic Impact Analysis: Methodology and Applications. Edited by Soul Pleeter. Boston, MA, Martinus Nijhoff Publishing Company 1980, pp. 85-113.

Doeksen, Gerald and Little, Charles. "Effect of Size of Input-Output Model on the Results of an Impact Analysis," *Agricultural Economic Research*, 1968, volume 20, number 4, pp.134-138.

Eidel'man, M. "The New Ex-Post Interbranch Balance of Production and Distribution of Output in the National Economy of the USSR," *Problems of Economics*, volume 16, number 1, May 1973, pp. 3-34.

Eidel'man, M. "National Economic Proportions According to Data of the Interbranch Balance," *Problems of Economics*, February 1970, pp. 28-50.

Eidel'man, M. "Problems in Constructing a New Ex-Post Interbranch Balance of the USSR National Economy for 1966," *Problems of Economics*, volume 11, number 3, July 1968, pp. 17-32.

Emerson, M. Jarvin. "Interregional Trade Effects in Static and Dynamic Input-Output Models." in Polenski, Karen and Skolka, Jiri, (eds.), *Advances in Input-output Analysis*: Proceedings of the Sixth International Conference on Input-Output Techniques, Vienna, April 1974. Lippincott, Balinger, Cambridge, MA, 1976, pp. 263-277.

Evans, W. D. "The Effect of Structural Matrix Errors on Interindustry Relations Estimates." *Econometrica*, volume 22, 1954, pp. 461-480.

Fei, J. C. H. "A Fundamental Theorem for the Aggregation Problem of Input-Output Analysis", *Econometrica*, volume 24, October 1956, pp. 400-412.

Fineburg, S. "An Iterative Procedure for Estimation in Contingency Tables." *Annals of Mathematics and Statistics*, volume 41, number 3, 1970, pp. 907-917.

Fisher, W. D. "Criteria for Aggregation in Input-Output Analysis," *Review of Economics and Statistics*, volume 40, August 1958, pp. 250-260.

Fontela, E., Dual,A., Gabus, A., Borlin, M., and Velay, C. "Forecasting technical coefficients and changes in relative prices." In *Applications of Input-Output Analysis*, Carter, A.P. and *Brody*, A. (editors), volume 2, Amsterdam, North-Holland Publishing Company, 1970, pp. 331-345.

Fox, L. *An Introduction to Numerical Linear Algebra*, New York: Oxford University Press, 1965.

Friedlander, D. "A Technique for Estimating a Contingency Table, Given the Marginal Totals and Some Supplementary Data." *Journal of Royal Statistical Society*, Series A, 123, Part 3, 1961, pp. 412-420.

Gallik, Dimitri M., Guill, Gene D., Kostinsky, Barry L., and Vladimir, G. "The 1972 Input-Output Table and the Changing Structure of the Soviet Economy in a Time of Change", in *Soviet Economy in a Time of Change*, a compendium of papers submitted to the Joint Economic Committee Congress of the United States, vol. I, Joint Committee Press, October 1979.

Gallik, Dimitri M., Heinmeier, Meredith., Kostinsky, Barry., Treml, Vladimir., Tretyakova, Albina. "*Construction of a 1977 Soviet Input-Output Table*." CIR Staff Paper U.S D.O.C Bureau of the Census, January 1984.

Gallik, Dimitri M., Kostinsky, Barry L, and Treml, V. *Input Output Structure of the Soviet Economy, 1972*, Foreign Economic Report Number 18 U.S. Department of Commerce, Bureau of the Census, Washington, D.C., 1983.

Garhart, Robert E., and Giarratani, Frank "Nonsurvey Input-Output Estimation Techniques: Evidence on the Structure of Errors," *Journal of Regional Science*, 1987, volume 27 number 2, pp. 245-253.

Geary, R. C. "A Method of Estimating the Elements of an Interindustry Matrix Knowing the Row and Column Tables," *Economics and Social Review*, volume 4, number 4, July 1973, pp. 477-485.

Ghosh, A. *Experiments with Input-Output Models: An Application to the Economy of the United Kingdom, 1948-1955*, England: Cambridge University Press, 1964.

Gerkin, Shelby D. "*Estimation of Stochastic Input-Output Models.*" Leiden: Martinus Nijhoff, 1976.

Giarratani, Frank. "A note on the McMenamian-Haring Input-Output Projection Technique." *Journal of regional Science*, 15, number 3, December, 1975, pp.371-73.

Gibbons, Joel c. Wolsky, Alan, and Tolley, George. "Approximate Aggregation and Error in Input-Output Models," *Resources and Energy*, volume 4 number 3, 1982, pp. 203-230.

Glattfedder, P., and Vaczi, P. "*Some Remarks on Theory of the RAS method.*" Proceedings of Second Hungarian Conference on I-O Technique, Akademia Kiado, Budepest, 1976, pp. 293-302.

Gossling, W. F. *Medium-Term Dynamic Forecasting*, proceedings of the 1975 London Conference, London: Input-Output Publishing Company, 1977.

Gossling, W. F. *Capital Coefficients and Dynamic Input-Output Models*, London: Input-Output Publishing Company, 1975.

Gossling, W. F. ed., *Input-Output and Throughput*, proceedings of the 1971 Norwich Conference Input-Output Publication Company, London, 1975.

Grandville, O. dela, Fontela, E., and Gabus, A. "A Note on European Projections of Foreign Trade." *Economia Internazionale, Geneva*, volume 21, number 4, November, 1968, pp. 593-609.

Grassini, Maurizio, and Smyshliaev, Anatolii S. *Input-Output Modeling: Proceedings of the Third IIASA Task Force Meeting, 22-25 Sept. 1982*, Luxenburg (Austria): International Institute for Applied System Analysis, 1983.

Green, H.A. John *Aggregation in Economic Analysis*, 1964, Princeton, New Jersy, Princeton University Press.

Gretton, P. and Cotterel P. "*The RAS Method for Compiling Input Output Tables. Australian Bureau of Statistics' Experience.*" Eighth Conference of Economists. LA Trobe University pp. 27-31, August 1979.

Griffith, William E., Hill, R. Carter, Judge, George G. *Learning and Practicing Econometrics*, John Wiley & Sons, New York, 1993.

Harrigan, F., and Buchanan, I. "A Quadratic Programming Approach to Input-Output Estimation and Simulation," *Journal of Regional Science*, volume 24, number 3 (1984), pp.339-358.

Harrigan, F., McGilvray, J., and McNicoll, I. "Simulating the Structure of a Regional Economy," *Environment and Planning A*, volume 12 (1980). pp. 936-972.

Hartfiel, D. J. "Some Error Estimates in Matrix Economic Models", *Journal of Economic Theory*, volume 19, number 2, December 1978, pp. 551-554.

Hatfield, Mark C. *Structural Changes In the Soviet Economy, 1959 -1977: An Input Output Analysis*. Thesis, University of Illinois at Urbana-Champaign, 1987.

Hatanaka, M. "*A note on consolidation within a Leontief system.*" *Econometrica*, vol. 20, no. 2, 1952, pp.301-303

Henry, E. W. "Relative Efficiency of RAS Versus Least Squares Methods of Updating Input-Output Structures: An Addendum," *Economics and Social Review*, volume 5, number 2, January 1974, pp. 175-179.

Henry, E. W. "Relative Efficiency of RAS Versus Least Squares Methods of Updating Input-Output Structures as Adjusted by Application to Irish Data," *Economics and Social Review*, volume 5, No. 1, October 1973, pp. -7-29.

Hewings, G. J. D. "The Role of Prior Information in

Updating Regional Input-Output Models," *Socio Economic Planning Science*, Volume 18, 1984, pp. 319-336.

Hewings, Geoffrey, J.D. "Evaluating the Possibilities for Exchanging Regional Input-Output Coefficients," *Environment and Planning A*, volume 9 (1977), pp. 927-944.

Hewings, Geoffrey J.D. "Input-Output Models: Aggregation for Regional Impact Analysis," *Growth and Change*, volume 3, number 1, 1972, pp. 15-19.

Hewings, Geoffrey, J.D. "Regional Input-Output Models in the U.K.: Some Problems and Prospects for the Use of Non-Survey Techniques." *Regional Studies*, volume 5, number 1, 1971, pp.11-22.

Hewing, Geoffrey and Janson, B.N. "Exchanging Input-Output Coefficients: A Reply and Further Comments," *Environment and Planning A*. volume 12, number 7, July, 1980, pp. 843-854.

Hinojosa, R.C. " A Performance Test of The Biproportional Adjustment of Input-Output Coefficients." *Environment and Planning, A*, Volume 10, 1978, pp. 1047-1052.

Iaremenko, IU V., Lavrenov, N. A., and Sutiagin, V. S. "Calculation and Analysis of Interindustry Proportions in the National Economy of the USSR," *Matekon*, volume 11, number 4, Summer 1975, pp. 40-56.

Ijiri, Yuji. "Fundamental Queries in Aggregation Theory," *Journal of the American Statistical Association*, volume 66, 1971 pp. 766-782

Isserman, Andrew M. "The Location Quotient Approach to Estimating Regional Economic Impacts," *AIP Journal*, 1977, volume 44, pp.33-41.

Isaacson, E., and Keller, H.B. *Analysis of Numerical Methods*, New York, John Wiley and Sons, 1966.

Israilevich, Philip, Reuve. *Biproportional Forecasting of Input-Output Tables*, Ph.D dissertation, University of Pennsylvania, 1986.

Jensen, Rodney C. "The Concept of Accuracy in Regional Input-Output Models," *International Regional Science Review*, volume 5, number 2, Winter 1980, pp.139-54.

Jensen, Rodney C., and West, G.R. "The Effect of Relative

Coefficient Size on Input-Output Multipliers," *Environmental and Planning*, A, volume 12, 1980, pp. 659- 670.

Kaneko, Yukio. "*An Empirical Study on Projecting and Forecasting the Input-Output Coefficient Matrix in Leontief Model*." Discussion Paper number 2. Hiroshima University of Economics, December 1983.

Karaska, Gerald J. "Variation of Input-Output Coefficients for different Levels of Aggregation," *Journal of Regional Science*, volume 8, number 2, 1968, pp. 215-227.

Kendrick, David A. "On the Leontief Dynamic Inverse," *Quarterly Journal of Economics*, volume 86, number 4, November 1972, pp. 693-696.

Khan, Abdul Qayyum. *Development of Updating Procedures for and Analysis of Pakistan Interindustry Relationships*, Ph.D. dissertation, Colorado State University, 1988.

Kim, Keuck Soo. *Nonsurvey Input_Output Updating Techniques: A Comparison of Accuracy with Korean Data*, Ph.D. dissertation, University of Pittsburgh, 1984.

Klein, Lawrence, R. "On the interpretation of Professor Leontief's system." Review of Economic Studies, volume 20, 1953, pp.131-136.

Kossov, V. "The theory of aggregation in input-output models, chapter 12 of *Contributions to Input-Output Analysis*. Carter, A.P., and Brody, A. editors. North-Holland Publishing Company, Amsterdam, 1972.

Kostinsky, Barry L. "The Reconstructed 1966 Soviet Input-Output Table:Revised Purchasers' and Producers' Prices Tables." *Foreign Economic Report* number 13, U.S. Department of Commerce, Bureau of Economic Analysis, September 1976.

Kubo, Yuji, Robinson, Sherman, Urata, Shujiro. "The Impact of Alternative Development Strategies: Simulations with a Dynamic Input-Output Model." *Journal of Policy Modeling*, Winter 1986, volume 8, number 4, pp.503-29.

Kullback, S., and Leibler, R. "On Information and Sufficiency," *Annals of Mathematical Statistics*, volume 22, 1951, pp. 79-86.

Kymn, K. O. "Aggregation in Input-Output Models: A

Comprehensive Review." *Economic Systems Research*, 1990, volume 2, number 1, pp. 65-93.

Lahr, Michael L., and Stevens Benjamin H. "Sectoral Aggregation in Regional Input-output Models, Concepts and Problems", paper presented at the *North American Meetings of Regional Science Association*, November, 1987, Baltimore, Maryland.

Lamel, J., Richter, J., Taufelsbauer, W., and Zelle, K. "*Comparative Analysis of Mathematical Methods for Updating Input_Output Matrices*," paper presented at the Sixth International Conference on Input-output Techniques, Vienna, Austria, 1974.

Lee, Cheng F. *Statistics for Business and Financial Economics*, D.C. Heathand Company, Lexington, MA, 1993.

Lecomber, J. R. C. "*A Critique of Methods of Adjusting, Updating and Projecting Matrices*," chapter 1 of Allen, R.I.G. and Gossling, W.F. (eds.). *Estimating and Projecting Input-Output Coefficients*, London: Input-Output Publishing Company, 1975.

Lecomber, Richard. A *Critique of Methods of Adjusting, Updating and Projecting Matrices, Together with Some New Proposals*, Cambridge Growth Project, Discussion Paper in Economics, number 40, August 1971.

Lecomber, Richard. "RAS Projections When Two or More Complete Matrices are Known," *Economics of Planning*, volume 9, number 3, 1969, pp. 267-278.

Lecomber, J. R. C. "*A Generalization of RAS*," Cambridge, Department of Applied Economics, Growth Project Paper 196, 1964.

Leontief, Wassily. *Input-Output Economics*, Second Edition, New York: Oxford University Press, 1985. (first published in 1966).

Leontief, Wassily, Chenery, Hollis B., Clark, Paul G., Duesenberry, James S., Ferguson, Allen R., Grosse, Anne P., Grosse, Robert N., Holzman, Mathilda, Isard, Walter, and Kistin, Helen. *Studies in the Structure of the American Economy: Theoretical and Empirical Explorations in Input-Output Analyzing* White Plains (N.Y.): International Arts and Sciences Press, 1953, (1977 reprint.)

Leontief, Wassily. "Some Basic Problems of Structural

Analysis," *Review of Economic and Statistics*, volume 34, February 1952, pp. 1-9.

Leontief, Wassily. *The Structure of American Economy 1919-1939*, 2nd edition. IASP publishing Company New York, 1951. (first published in 1941)

Leontief, Wassily. "Recent Development in the Study of Interindustrial Relationships," *American Economic Review*, volume 39, May 1949, pp. 211-225.

Leontief, Wassily. "Structural Matrices for National Economies," *Econometrica*, volume 17, supplement July 1949, pp. 273-280.

Leontief, Wassily. "Quantitative Input and Output Relations in the Economic System of the United States," *Review of Economic and Statistics*, volume 18, August 1936, pp. 105- 125.

Long, N.B. "An Input-Output Comparison of the Economic Structure of the US and USSR", *Review of Economic and Statistics*, volume 52 number 4, 1970.

Lu, Alan Yun. *Estimation of Potential GNP in the U.S. Economy: An Application of I-O Analysis and Alternative Updating Methods*, Nankang, Taipei, Taiwan, Republic of China: Institute of Economics, Academia Sinica, 1983.

Lynch, R.G. "An Assessment of the RAS Method for Updating Input-Output Tables," In: *Readings In Input-Output Analysis*, Ira Sohn, (ed.), New York, Oxford University Press, 1986, pp. 271-284.

Macgill, Sally M. "Convergence and Related Properties for a Modified Biproportional Matrix Problem." *Environment and Planning, A*, 11(1979), pp 499-506.

Macgill, Sally M. "Theoretical Properties of Biproportional Matrix Adjustment." *Environment and Planning A*, volume 9, (1977), pp. 687-701.

Malizia, Emil, and Bond, Daniel L. "Empirical Tests of the RAS Method of Interindustry Coefficient Adjustment," *Journal of Regional Science*, volume 14, number 3, December 1974, pp. 355-365.

Matuszewski, T. I., Pitts, P. R., and Sawyer, J. A. "Linear Programming Estimates of Changes in Input-output Coefficients," *Canadian Journal of Economics*, volume 30, May 1964, pp. 203-210.

McGilvray, James W., and Simpson, David. "Some Tests of Stability in Interindustry Coefficients", *Econometrica*, volume 37, number 2, April 1969, pp. 204-221.

McMenamin, David G., Haring, Joseph E. "An Appraisal of Non-survey Techniques for Estimating Regional Input-Output Models." *Journal of Regional Science*, August 1974 volume 14 number 2, P.P.191-205.

Miernyk, William H. "The Projection of Technical Coefficients for Medium Term Forecasting" in W. F. Goosling, (ed.), *Medium Term Dynamic Forecasting*, the 1975 London Input-Output Conference. London, Input-Output Publishing Company, 1977, pp.29-41.

Miernyk, William H. "Comments on Recent Developments in Regional Input-Output Analysis," *International Regional Science Review*, Volume 1, number 2, 1976, pp. 47-55.

Miernyk, William H. *The Elements of Input-Output Analysis*, Random House Publishing Company, New York, 1965.

Miller, Ronald E., and Blair, Peter. *Input-Output Analysis: Foundation and Extensions*, Englewood Cliffs (N.J.), Prentice Hall, 1985.

Mohr, Malte. *A User's Guide to Constrained Biproportional Matrix Estimation*. DOT Report # TST-76-62, September 1975.

Mohr, Malte, Crown, William H., and Polenske, Karen R. "A Linear Approach to Solving Infeasible RAS Problems." *Journal of Regional Science*, Volume 27, number 4 1987, pp. 587-603.

Morimoto, Yoshinori. "A Note on Weighted Aggregation in Input-Output Analysis," *International Economic Review*, volume 12, number 1, February 1971, pp. 138-143.

Morimoto, Yoshinori. "On Aggregation Problems in Input-Output Analysis," *Review of Economic Studies*, volume 37, number 1, January 1970, pp. 119-126.

Morishima, M. *Equilibrium, Stability and Growth; A Multi Sectional Analysis*, Clarendon Press, Oxford, 1964.

Morishima, M. A. "A comment on Dr. Klein's interpretation of Leontief's system," *Review of Economic Studies*, volume 24, 1957, pp. 65-68.

Morrison, W.I. and Smith, P. "Nonsurvey Input-Output

Techniques at the Small Area Level: An Evaluation," *Journal of Regional Science*, volume 14, February, 1974, pp.1-14.

Morrison, W.I. and R.G. Thuman. "A Lagrangian Multiplier Approach to the Solution of a Special Constrained Matrix Problem." *Journal of Regional Science*, volume 20 (1980), pp 279-292.

Noble, B., *Applied Linear Algebra*, Prentice-Hall, Inc., 1969.

Omar, F.H. *The Projection of Input-Output Coefficients with Application to the United Kingdom*, Ph.D. Thesis, University of Nttingham, 1967.

Parikh, Ashok. "Forecasts of Input-Output Matrices Using the RAS Method", *Review of Economics and Statistics*, volume 61, number 3, August 1979, pp. 477-481.

Pearson, Carl E. *Handbook of Applied Mathematics: Selected Results and Methods*, New York, Van Nastrand Reinhold Company, 1983.

Pigozzi, Bruce W., and Hinojosa, Rene C. "Regional Input-Output Inverse Coefficients Adjusted from National Tables." *Growth and Change*, volume 16, January 1985, pp.8-12.

Pitzer, John. *An Analysis of Technical Changes in the Soviet Economy: An Application of Soviet Input-Output Tables*, American University, 1980

Polenske, Karen R. "*Analysis of Output, Income, and Employment Impacts in Regional and Multiregional Models*." Report number 31. Paper presented at the Session in Honor of William H. Miernyk: Interregional and Multiregional Models at the Twenty-ninth North American Meetings of the Regional Science Association, Pittsburgh, P.A, November 12, 1982.

Polenske, Karen R., Crown, William H., and Mohr, Malte. *A Critical Review of the RAS Literature*. Report # 36. Presented at the Strategic Regional Policy Workshop, Warsaw, Poland, December 12, 1984 and the Second Soviet American Seminar on Regional Planning, Tillin, Estonia, U.S.S.R., January 7, 1986

Polenske, Karen R., and Skolka, Jiri V., eds. *Advances in Input Output Analysis: Proceedings of the Sixth International Conference on Input-Output Techniques*, *Vienna, 22-26 April 1974*, Cambridge (Mass.), Lippincott, Ballinger, 1976.

Ralston, Scott N., Hastings, Steven E., and Brucker, Sharon M. "Improving Regional I-O Models: Evidence Against Uniform Regional Purchase Coefficients," *Annals of Regional Science*, volume 20, number 1, 1986, pp. 65-80.

Richardson, Harry W. *Input-Output and Regional Economics*, Halsted Press, New York, 1972.

Round, Jeffrey I. "Nonsurvey Techniques: A Critical Review of the Theory and Evidence." *International Regional Science Review*, December 1983, volume 8, number 3, pp. 189- 212.

Round, Jeffrey, I. " An Interregional Input-Output Approach to the Evaluation of Non-Survey Methods." *Journal of Regional Science*, 18, number 2, August 1978, pp.179-94.

Sawyer, C. H., (1992)

Sawyer, C.H. and Miller, R.E. "Experiments in Regionalization of a National Input-Output Table," *Environment and Planning A*, volume 15, 1983, pp. 1501-1520.

Sevaldson, P. "The Stability of Input-Output Coefficients." pp 207-237 of *Application of Input-Output Analysis*, edited by Carter, A. P. and Brody, A., 1970, Amsterdam, North-Holland.

Schaffer, Norman C. *An Analysis of the Assumption and Updating Procedures of National Input-Output Tables Used in Economic Forecasting and Planning*, Clemson University, 1970.

Schaffer, William A. and Chu Kong. "Nonsurvey Techniques for Constructing Regional Interindustry Models," *Papers and Proceedings of the Regional Science Association Meetings*" volume 23 (November) 1969.

Schneider, Howard M. "*An Evaluation of Two Alternative Methods for Updating Input-Output Tables*," B.A Dissertation, Harvard University, Cambridge, MA, 1965.

Sebald, Anthony V. *An Analysis of the Sensitivity of Large Scale Input-Output Models to Parametric Uncertainty*, Thesis, University of Illinois at Urbana-Champaign, 1974.

Sherman, J., and Morrison W. "Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix," *Annals of Mathematical Statistics*, volume 21, 1950, pp. 124-127

Simonovits, A. "A Note on the Underestimation and Overestimation of the Leontief Inverse," *Econometrica*, volume 43, number 3, May 1975, pp. 493-98

Simpson, D. and Tuskui, J. "The Fundamental Structure of Input-Output Tables, an International Comparison," *Review of Economics and Statistics*, volume 47, 1965, pp. 434-446.

Slater Lucy Joan. "*A General Econometric Matrix Program.*" University of Cambridge, Department of Applied Economics Occasional Paper # 46, 1976.

Smyshlyaev, Anatoli, ed. *Input-Output Modeling: Proceedings of the Fifth IIASA (International Institute for Applied Systems Analysis) Task Force Meeting on Input-Output Modeling held at Laxenburg, Austria, 4-6 Oct. 1984*, Berlin, New York, Springer-Verlag, 1985.

Smyshlyaev, Anatoli, ed. *Proceedings of the Fourth Task Force Meeting on Input-Output Modeling, 29 September-1 October 1983*, Laxenburg (Austria), International Institute for Applied Systems Analysis, 1983.

Snickars, F., and Weibul, J. "A Minimum Information Principle," *Regional Science and Urban Economics*, volume 7, 1977, pp. 137-168.

Snower, Daniel J. "New Methods of Updating Input-Output Matrices," *Economic Systems Research*, volume 2 number 1, 1990, pp.27-37.

Sohn, Ira (ed.) *Readings in Input-Output Analysis: Theory and Applications*. New York: Oxford University Press 1986.

Sonis, Michael, and Hewings, Geoffrey, J. D. "*Error and Sensitivity Input-Output Analysis: A New Approach,*" Frontiers of Input-Output Analysis, New York: Oxford University Press, 1989.

Spulber, Nicholas, and Moayed-Dadkhah, Kamran. "The Pioneering Stage in Input-Output Economics: The Soviet National Economic Balance 1923-24, After Fifty Years," *Review of Economics and Statistics*, volume 57, number 1, February 1975, pp. 27-34.

Statistical Publishing House. *Input-Output Techniques* Proceedings of the Third Hungarian Conference on Input-Output Techniques, 3-5 November 1981, Budapest 1982.

Stephan, F.F. "An Iterative Method of Adjusting Sample Frequency Tables when Expected Marginal Totals are Known", *Annals of Mathematical Statistics*. volume 13, 1942

St.Louis, Larry V. "Empirical Tests of Some Semi-Survey Update Procedures Applied to Rectangular I-O Tables," *Journal of Regional Science* volume 29, number 3, 1989, pp.373-85.

Stone, Richard, Bates, John, and Bacharach, Michael. *A Program for Growth*, Input-Output Relationships, 1954-1966, volume 3, Cambridge, MA, the MIT Press,1963.

Stone, Richard, and Brown, A. "A Computable Model of Economic Growth," volume one of *A Program for Growth* Series, London: Chapman and Hall, 1962.

Stone, Richard. *Input-Output and National Accounts*, OEEC, Paris, 1961.

Strang, Gilbert. *Linear Algebra and Its Applications*, Second Edition, Academic Press, New York, 1980.

Teibout, C. M. "An Empirical Regional Input-Output Projection Model," *Review of Economics and Statistics*, volume 51, 1969, pp. 334-340.

Theil, H. *Economics and Information Theory*, North Holland, Amsterdam, 1967.

Theil, H. *Applied Economic Forecasting*, London, North-Holland Publishing Company, 1966 .

Theil, H. "Linear Aggregation in Input-Output Analysis," *Econometrica*, volume 25, January 1957, 111-122.

Tilanus, C. B. "Marginal Versus Average Input Coefficients in Input-Output Forecasting," *Quarterly Journal of Economics*, volume 81, February 1967, pp. 140-145 .

Tilanus, C. B. *Input-Output Experiments: The Netherlands 1948 1961*, Rotterdam University Press, 1966.

Tilanus, C.B. and G. Rey. "Input-Output Volume and Value Predictions for the Netherlands, 1948-1958."

International Economic Review, volume 5, number 1, January 1964, pp. 34-45.

Treml, Vladimir G.(ed.) *Studies in Soviet Input-Output Analysis*. New York, Praeger Publishers, Inc. 1977.

Treml, Vladimir G. "Characteristics of the U.S.S.R. Input-Output Tables." In: International Comparisons of Interindustry Data. *United Nations Industrial Planning and Programming Series*, number 2, P.P. 133- 153.

Treml, Vladimir G. *Input-Output Analysis and the Soviet Economy: An Annotated Bibliography*, New York: Praeger,1975.

Treml, Vladimir G. "A Note on Soviet Input-Output Tables",*Soviet Studies*, volume 21, number 1, July 1969, pp. 21-34.

Treml, Vladimir G., Gallik, Dimitri M., Kostinsky, Barry L., Kruger, Kurt W. *The Structure of the Soviet Economy: Analysis and Reconstruction of the 1966 Input-Output Table*. New York, Praeger Publishers Inc., 1972.

Treml, Vladimir G., Gallik, Dimitri M., Kostinsky, Barry L., Kurtzweg, Laurie R., and Tretyakova, Albania F. The Soviet 1966 and 1972 Input-Output Tables." Joint Economic Committee, U.S. Congress. *Soviet Economy in a New Perspective*. Washington, D.C., 1976.

Treml, Vladimir G., Kostinsky, Barry L., and Gallik, Dimitri M. "*Interindustry Structure of the Soviet Economy: 1959 and 1966*".

Treml, Vladimir G., Kostinsky, Barry L., Kruger, Kurt W., and Gallik, Dimitri, M. Conversion of Soviet Input-Output Tables to Producers' Prices: the 1966 Reconstructed Table. U.S. Department of Commerce, *Foreign Economic Reports*, number 1, Washington, D.C., 1973.

Tsukui, Jinkichi, Hori, Hajime. "Rolling Planes and the Turnpike in a Dynamic Input-Output System: A Simulation Study." *Journal of Policy Modeling*, Spring 1989, volume 11, number 1, P.P. 91-109.

United Nations. *Input-Output Models: Theory, Data and Application*, United Nations Economic and Social Commission for Asia and the Pacific; Statistical Division, Tokyo, Japan: Institute of Developing Economies, 1985.

United Nations Industrial Development Organization. *Proceedings of the Seventh International Conference on Input-Output Techniques*, New York, United Nations, 1984.

Uribe, Pedro. "Some 'RAS' Experiments with the Mexican Input-Output Model," *Annals of Economic and Social Measurement*, volume 4, number 4, 1975, pp. 553-569.

Uribe, P., deLeeuw, C., and Theil, H. "*The Information Approach to the Prediction of Interregional Trade Flows*," Report number 650, Econometric Institute of Netherlands, School of Economics, Rotterdam, 1965.

Vaccara, Beatrice N. "*Changes Overtime in Input-Output Coefficients for the United States*." Office of Business Economics U.S. Department of Commerce, 1969

Viet, Vu Quang. *Sensitivity Analysis in Input-Output: Theory and Application*, New York University, 1980.

Volodarskii, L., and Eidel'man, Michio. "Main Results of the Ex-Post Interbranch Balance for 1972", *Problems of Economics*, volume 18, number 11, March 1976, pp. 54-74.

Volodarskii, L., and Eidel'man, Michio. "Basic Results of Elaboration of the Interbranch Balance of Production and Distribution of Output Through the National Economy of the USSR for 1966", *Problems of Economics*, September 1969, pp. 29-51.

Watson, Collin J., Billingsley, Patrick, Croft, D. James, Huntsberger, David V. *Statistics For Management and Economics*, Fourth Edition, Allyn and Bacon, Boston, MA, 1990.

West, G.R. "Sensitivity and Key Sector Analysis in Input-Output Models." Australian Economic Papers, volume 21, 1982, pp. 365-378.

Williamson, Robert B. "Simple Input-Output Models for Area Analysis," *Land Economics*, volume 46, number 3, 1970, pp. 333-338.

Woslsky, A.M. "Disaggregating Input-Output Models." *Review of Economics and Statistics*, volume 66, number 2, May 1984, pp.283-291.

Yamada, Isamu. *Theory and Application of Interindustry Trade*, Kinokuniya Bookstore Company, Tokyo, 1961.