INFORMATION TO USERS

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

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

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
TANG, LE DAk

A CAUSAL ANALYSIS OF DEVELOPMENT OF TRADE AND INDUSTRIAL
NEW AND EMERGING OCCUPATIONS IN MANUFACTURING

The Ohio State University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106
A CAUSAL ANALYSIS OF DEVELOPMENT OF TRADE AND INDUSTRIAL
NEW AND EMERGING OCCUPATIONS IN MANUFACTURING

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Le Dak Tang, B.Ed., M.A.

*****

The Ohio State University
1981

Reading Committee:
Dr. Dewey A. Adams
Dr. Aaron J. Miller
Dr. Thomas N. Chirikos

Approved By

Dewey A. Adams
Adviser
Academic Faculty of
Vocational-Technical
Education
Dedicated

To My Parents

S. F. Tang and P. H. Tang
ACKNOWLEDGMENTS

My sincere appreciation is extended to the members of my reading committee. Dr. Dewey A. Adams, Chairman, offered continuous support, counsel, and encouragement, and intellectual stimulation during this study and throughout my doctoral program. Dr. Thomas N. Chirikos led me to the field of Human Resources Planning and provided me with productive critiques and suggestions on my dissertation. Dr. Aaron J. Miller helped guide the scope of this research and provided continuous support for my dissertation. All members of the committee have been sources of inspiration during my studies at The Ohio State University.

Gratitude is extended to Dr. H. Lawrence Hotchkiss, Research Specialist at the National Center for Research in Vocational Education, whose laborious discussions on issues in social science research methodology helped stimulate a broader perspective for the methods used in this study. Also included are Mark Shaff, Director of the Ohio Occupational Information Coordinating Committee, and various individuals in the Bureau of Labor Statistics, United States Department of Labor, who provided me with published as well as unpublished data pertinent to the study.

Finally, and most importantly, I express appreciation to my wife, Karen, who edited my dissertation as well as provided constant support, encouragement, and patience in bringing this study to its completion.
VITA

October 22, 1949 . . . Born in Kwantong, China

1972 . . . . . . B.Ed. in Industrial Education, National Taiwan Normal University, Taipei, Taiwan, Republic of China

1972-1973 . . . . Vocational Teacher, Pu Le Vocational-Industrial School, Pu Le, Taiwan, Republic of China

1973-1975 . . . . Military Service

1975-1976 . . . . Vocational Teacher, Taoyuan Vocational-Industrial School, Taoyuan, Taiwan, Republic of China

1976-1977 . . . . Teaching Assistant, Electronics Lab, East Tennessee State University, Johnson City, Tennessee

1977 . . . . . . M.A. in Industrial Technology, East Tennessee State University, Johnson City, Tennessee

1977 . . . . . . Research Assistant, Department of Industrial Education, Texas' A & M University College Station, Texas

1978-1981 . . . . Graduate Research Associate, National Center for Research in Vocational Education, The Ohio State University, Columbus, Ohio

PUBLICATIONS


FIELDS OF STUDY

Major Field: Vocational-Technical Education

Studies in Vocational-Technical Education.
Professor Dewey A. Adams, Professor Aaron J. Miller

Studies in Human Resources Planning.
Professor Thomas N. Chirikos
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>VITA</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>x</td>
</tr>
</tbody>
</table>

**Chapter**

I. INTRODUCTION ........................................ 1
   - Background and Statement of the Problem ........ 1
   - Purposes of the Study ................................ 2
   - Significance of the Study ........................... 3
   - The Scope of the Study ................................ 5
   - Procedures of the Study .............................. 6
   - Definitions ........................................... 9
   - Limitations ........................................... 10
   - Assumptions ........................................... 10
   - Summary ............................................... 11

II. SUBSTANTIVE CONSIDERATIONS ......................... 12
   - What Are New and Emerging Occupations? .......... 12
   - Economic Forces for Development of New and Emerging Occupations ...................... 15
   - Technological Change ................................ 15
   - Labor Market Process ................................ 17
   - Building a Model ..................................... 20
   - Variables and the Model ............................. 20
   - The Hypotheses ....................................... 27
   - From Conceptual Model to Multiple Indicators
     - Model ............................................... 28
     - Measurement of the Variables ...................... 29
     - The Multiple Indicators Model .................... 38
   - Summary ............................................... 39
### III. DATA CONSIDERATIONS

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population and Sample</td>
<td>43</td>
</tr>
<tr>
<td>Selection of Substitutable Occupations</td>
<td>47</td>
</tr>
<tr>
<td>Data for Endogenous Variables</td>
<td>48</td>
</tr>
<tr>
<td>Data for New and Emerging Occupations</td>
<td>48</td>
</tr>
<tr>
<td>Data for Occupational Demand Effect</td>
<td>50</td>
</tr>
<tr>
<td>Data for Occupational Supply Effect</td>
<td>53</td>
</tr>
<tr>
<td>Data for Occupational Structure Change</td>
<td>55</td>
</tr>
<tr>
<td>Data for Sectoral Growth</td>
<td>57</td>
</tr>
<tr>
<td>Data for Technological Change</td>
<td>58</td>
</tr>
<tr>
<td>Data for Exogenous Variables</td>
<td>60</td>
</tr>
<tr>
<td>Data for Research and Development</td>
<td>60</td>
</tr>
<tr>
<td>Summary</td>
<td>62</td>
</tr>
</tbody>
</table>

### IV. METHODOLOGICAL CONSIDERATIONS

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specification</td>
<td>65</td>
</tr>
<tr>
<td>Identification</td>
<td>74</td>
</tr>
<tr>
<td>Estimation</td>
<td>77</td>
</tr>
<tr>
<td>Model Testing</td>
<td>87</td>
</tr>
<tr>
<td>Summary</td>
<td>90</td>
</tr>
</tbody>
</table>

### V. MODEL TESTING

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter Estimates Using LISREL</td>
<td>96</td>
</tr>
<tr>
<td>Interpretations of the Results</td>
<td>104</td>
</tr>
<tr>
<td>Summary</td>
<td>111</td>
</tr>
</tbody>
</table>

### VI. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>113</td>
</tr>
<tr>
<td>Conclusions</td>
<td>115</td>
</tr>
<tr>
<td>Recommendations</td>
<td>121</td>
</tr>
</tbody>
</table>

### APPENDIX: TRADE AND INDUSTRIAL NEW AND EMERGING OCCUPATIONS IN MANUFACTURING

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIBLIOGRAPHY</td>
<td>128</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Exogenous and Endogenous Variables in the Model</td>
<td>40</td>
</tr>
<tr>
<td>2.</td>
<td>Means and Standard Deviations of the Indicators</td>
<td>92</td>
</tr>
<tr>
<td>3.</td>
<td>Correlation Coefficients of the Indicators</td>
<td>93</td>
</tr>
<tr>
<td>4.</td>
<td>Multiple Regression Using CNEO as Dependent Variable</td>
<td>94</td>
</tr>
<tr>
<td>5.</td>
<td>Significant Presence of the Parameters</td>
<td>95</td>
</tr>
<tr>
<td>7.</td>
<td>Model Fitting Allowing for Correlated Measurement Errors</td>
<td>100</td>
</tr>
<tr>
<td>8.</td>
<td>Reliability Coefficients in Successive Measurement Models</td>
<td>100</td>
</tr>
<tr>
<td>9.</td>
<td>Parameter Estimates for Model B1</td>
<td>103</td>
</tr>
<tr>
<td>10.</td>
<td>Incremental Fit of Model Comparisons</td>
<td>104</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Flowchart of the Procedures</td>
<td>7</td>
</tr>
<tr>
<td>2.</td>
<td>A Conceptual Model for Development of T &amp; I New and Emerging Occupations</td>
<td>26</td>
</tr>
<tr>
<td>3.</td>
<td>A Multiple Indicators Model for Development of T &amp; I New and Emerging Occupations</td>
<td>41</td>
</tr>
<tr>
<td>4.</td>
<td>Multiple Indicators Causal Model with Symbols and Error Terms</td>
<td>72</td>
</tr>
</tbody>
</table>
A CAUSAL ANALYSIS OF DEVELOPMENT OF TRADE AND INDUSTRIAL NEW AND EMERGING OCCUPATIONS IN MANUFACTURING

By

Le Dak Tang, Ph.D.

The Ohio State University, 1981

Professor Dewey A. Adams, Adviser

The purposes of this study were (1) to develop a causal model to explain the structure underlying the developmental behaviors of Trade and Industrial new and emerging occupations in manufacturing and (2) to verify empirically the model developed and make necessary refinements. There were two major economic forces, technological change and labor market process, by which the development of new and emerging occupations was suggested. A conceptual model was proposed based on these economic forces. The variables included in the model were selected based upon the theories, empirical evidence, and postulations gathered from the literature review. A multiple indicators model was then developed, since most of the factors identified were latent variables. The data were from published as well as unpublished sources. The canned program, LISREL, was used in data analysis.

The goodness of fit of the model based on statistical (chi-square) and practical (Incremental Fit Index) criteria was well accepted. The reliability coefficients of the indicators representing
the latent variables were in reasonable range. The parameter estimates representing the major hypotheses of impact of occupational structure change, demand effects, and supply effects on changes in distributions of T & I new and emerging occupations in manufacturing were not statistically significant. However, their magnitudes were not too far from achieving significance.

In general, the causal development of T & I new and emerging occupations in manufacturing followed more closely to the economic force of technological change than it did to labor market process of demand and supply interactions. However, owing to the problems in data and some non-random measurement errors in the model, the results should be considered directive instead of conclusive.
Chapter I

INTRODUCTION

1.1 BACKGROUND AND STATEMENT OF THE PROBLEM

Although planning for vocational education in the United States has its historical roots in the period prior to the enactment of the Smith-Hughes Act in 1917, it is only since the 1968 and 1976 Amendments to the Vocational Education Act that such processes have been mandated in each state. This legislation has, among other things, charged vocational educators with the responsibility of providing curriculum design and development for new and emerging job fields or careers in order to meet changing labor market conditions. Since the original legislation was not designed with new and emerging occupations in mind, several questions concerning the intent of the legislation in this regard have been raised. One relates to the meanings of new and emerging occupations, because they were neither defined in the law nor in the rules and regulations. Given that they were not defined, another question is how do vocational educators determine what occupations are new and emerging and how are they identified? A third question is how do vocational educators determine needs for new and emerging occupations in order to justify curricula development activities? In other words, what are the demand and supply
situations of those new and emerging occupations which are appropriate for vocational education and training? A fourth question is how much do vocational educators know about the developmental behavior of those so-called new and emerging occupations?

If only curriculum design and development were involved, many planning models could be used to answer these questions. However, neither sufficient knowledge about the external factors influencing the curriculum needs nor relevant data on the needs were available. Furthermore, little knowledge has been accumulated about the process by which new occupations are being developed, the factors influencing their development, and the interrelationships among the factors.

It seemed appropriate at this stage to make an inquiry about the basic structure underlying the behavior of development of new and emerging occupations, especially for those appropriate for vocational training. This inquiry could bring additional information into the planning process from which actions in connection with new and emerging occupations could be guided.

1.2 PURPOSES OF THE STUDY

The major purpose of this study was to provide vocational educators and planners with basic understanding about the factors influencing the development of new and emerging occupations appropriate for Trade and Industrial (hereinafter referred to as T & I) training in manufacturing. Within this frame of inquiry, the study focused on the following questions. First, is it possible to
formulate a conceptual model for development of new and emerging occupations in manufacturing? If so, what structure should this model have and what determining factors should be included in it? Then, does this model lend itself to empirical verification? Lastly, how can this model be used to help vocational educators plan their programs in response to the demand for qualified persons for new and emerging occupations? Constructing a model to explain the causal relationships among the factors influencing the developmental behavior and empirically verifying the model were the first steps toward the answer. Specific purposes to be achieved by this study were:

(1) The development of a model to explain the causal structure underlying the developmental behaviors of T & I new and emerging occupations in manufacturing.

(2) The empirical verification of the model developed and the making of necessary refinements.

1.3 SIGNIFICANCE OF THE STUDY

Vocational education is vital to economic development of the nation in terms of its contribution to the quality of the labor force. As the economy progresses, a corresponding change in labor force in which the change in occupational structure is the major determinant is necessary. Vocational education can facilitate this changing process by adjusting its program offerings to the needs of the changing occupational structure. Then, a responsive labor force can be in place, and consequently, enhance the economic development.
For the preceding reason, vocational education is charged with the responsibility to provide relevant education and training for both the needs of students and the needs of the labor market. Planning vocational education in advance for structural change in occupations is to make sure that the labor supply will match the demand in the future. And development of new and emerging occupations is certainly part of the specific occupational change. Planning programs for new and emerging occupations is essentially to bridge the needs between labor market and students. However, the lack of knowledge about the developmental behavior of new and emerging occupations presents an impediment to any effective planning.

There has been little knowledge accumulated to further understanding of the behavior and the contributing factors to the development of new and emerging occupations in general, and development of new and emerging occupations appropriate for vocational training in specific. In order to undertake the task of planning vocational programs for new and emerging occupations in response to changing labor market conditions, vocational educators should acquire themselves the basic knowledge underlying this development. It appeared that the macro-picture of their developmental behavior, the contributing factors and their interrelationships, and the relative importance of these factors to the development could form the foundation of this basic knowledge. This knowledge in itself might not be sufficient to "cover" the entire planning effort. Nevertheless,
it would provide vocational educators with a common, and hopefully, a sound forum within which information exchange in this particular regard could be enhanced.

The intention of this study was exactly to fill this knowledge gap that was perceived as the impediment for vocational educators and planners to go further in dealing with new and emerging occupations in their program planning efforts. The study provides a basic foundation for understanding the developmental behavior of new and emerging occupations, especially for those in T & I area. Vocational educators and planners will be able to know the forces behind the changes in distribution of T & I new and emerging occupations and the policy implications if they try to influence the forces one way or the other. In addition, this subsequently provides a basis for further knowledge generation in connection with planning vocational education for new and emerging occupations.

1.4 THE SCOPE OF THE STUDY

There are five traditional service areas (agriculture, distribution and marketing, home economics, business and office, and trade and industry) in vocational education. Each area has its own historical background and characteristics. It was perceived that the basic structure governing the developmental behavior of new and emerging occupations in each service area might be similar but would not be exactly the same. Moreso would be the degree of contribution of each factor to the structure. Thus, on the one hand, trying to
establish an explanatory structure which could apply to all service areas might be very difficult, if not impossible. And, on the other hand, trying to undertake a study which could provide separate explanations for each of the service areas was equally difficult. So, a compromise needed to be made in terms of the scope of the study.

Based on the preliminary investigation about the distribution of new and emerging occupations in different service areas, it was found that the number of new and emerging occupations in Trade and Industry (T & I) area were far greater than those in other areas. In addition, almost all of the T & I occupations were distributed in manufacturing, which makes data collection easier and the results more interpretable. Therefore, the developmental behavior of only new and emerging occupations in T & I area was studied.

1.5 PROCEDURES OF THE STUDY

The research procedures can be followed conceptually in the sequences outlined in the flowchart in Figure 1.

The procedures were explained in detail as follows:

1. Literature review was to survey the existing body of knowledge surrounding the subject, particularly theoretical and empirical studies done in the area. The review led to a conceptual framework and subsequently to the considerations in theories, methods, and data.

2. Conceptual framework was to build a conceptual paradigm underlying the development of new and emerging occupations.
Figure 1: Flowchart of the Procedures
3. There were three dimensions: substantive, methodological, and data, which were to be considered simultaneously in the study. In substantive considerations, several things were of concern. First, the theories related to the plausible explanations of the behaviors of interest needed to be surveyed and organized. Second, theoretical constructs related to the phenomenon and their interrelationships needed to be identified. Third, operationalization or measurement of those theoretical constructs needs to be examined.

4. In methodological considerations, the following were of concern. Different schemes of model construction, different methods of parameters estimation and their implications to the results and interpretations of the model, different methods in model testing, modification, and the measurement errors needed to be studied.

5. In data considerations, the population, sample, sampling scheme, sources of data, availability of data, transformation of data, and ways of data collection and analysis, were considered.

6. Hypotheses concerning the causal orderings of the theoretical constructs were generated according to the theories and evidence gathered and were incorporated into the model.

7. The model was formulated based on the interactions of theories (hypotheses), method, and data. In this stage, variables in the model were identified and operationalized.

8. Model testing and modification were two feedback processes. Essentially, they were in so-called data analysis stage. The purpose was to obtain the model that could fit the data well in conjunction with prior substantive considerations.
9. The results were interpreted based on the hypotheses tested and their alternative explanations.

10. Conclusions were drawn based on the results and their interpretations. Recommendations for further research and implications for practice in planning vocational education for new and emerging occupations were made according to the research findings and existing knowledge.

1.6 DEFINITIONS

1. Causal analysis: The analysis of cause-and-effect relationships among the variables representing a phenomenon under non-experimental conditions.

2. Cause indicator: An indicator which is a cause for the theoretical construct that is measured by the indicator.

3. Effect indicator: An indicator which is an effect of the theoretical construct that is measured by the indicator.

4. Endogenous variable: The value of the variable is to be determined by the phenomenon expressed in the model. On some occasions, it is called dependent variable.

5. Exogenous variable: The value of the variable is determined independently or outside of the phenomenon. On some occasions, it is called independent or predetermined variable.

6. Latent variable: Variable that is not directly observable.

8. Occupational crosswalk: Directory of equivalent codes among different occupational classification systems.

9. Trade and Industry: One traditional service area out of five (agriculture, distribution and marketing, home economics, business and office, and trade and industry) in vocational education.

1.7 LIMITATIONS

There were four limitations in this study.

1. The study was limited to socio-economic variables in explaining the causal relationships of development of T & I new and emerging occupations in manufacturing in the United States during the period of 1965 to 1977. Other variables—for instance, psychological variables—were not considered in the study.

2. The model developed was limited to the theories and empirical evidence surveyed.

3. The sources of data were limited to the published information. Hence, the availability of data was limited to the existing sources.

4. The reliability of the data was dependent upon the methods and procedures employed by the agencies collecting the data.

1.8 ASSUMPTIONS

Six assumptions were made in the study.

1. It was assumed that causal orderings could be established among the variables affecting the development of T & I new and emerging occupations in manufacturing.
2. It was assumed that causal relationships among the variables were linear. Therefore, a structural equation model derived from a general linear model was appropriate to this study.

3. It was assumed, contrary to the convention, that in the causal model, the variables would be measured with errors. Therefore, measurement errors would be treated explicitly.

4. The assumptions of a linear structural model, for example, zeros means of errors, constant variance of the variables, uncorrelated variables and errors, and normally distributed error terms, must also hold asymptotically (in large sample size) in the study.

5. It was assumed that hypothesized relationships in the model adequately reflect current status of knowledge being surveyed.

6. It was assumed that temporal equilibrium of the developmental behavior has been achieved for those T & I new and emerging occupations identified in the population.

1.9 SUMMARY

This chapter included eight sections. The background of the study, the statement of the problem, the purposes, the significance, the scope, and the procedures of the study were treated in the first five sections. The definitions of the terms, the limitations of the study, and the assumptions of the study were described in the last three sections.
Chapter II

SUBSTANTIVE CONSIDERATIONS

A theoretical and conceptual model underlying the development of new and emerging occupations, the variables to be included in the model, and the procedures for their measurement were developed in this chapter. Initial consideration was given to a discussion of the meanings of new and emerging occupations. Then the major economic forces underlying the development of occupations were surveyed. The variables suggested by the theories were discussed next and a conceptual model was then presented. Then the measurement of those variables was discussed and a multiple indicators model incorporating both theoretical constructs and the measurement was presented.

2.1 WHAT ARE NEW AND EMERGING OCCUPATIONS?

Prior to a discussion of new and emerging occupations, it seemed appropriate to differentiate first of all between job and occupation, since they were used interchangeably throughout the study. Job referred to all the tasks carried out by a particular worker to complete his/her duties while occupation referred to a collection of jobs sufficiently similar in their main tasks to be grouped under a common title for classification purpose (British Department of Employment, 1972). In other words, occupations were identified and
grouped primarily in terms of the work usually performed, this being
determined by the tasks, duties, and responsibilities of the occu-
pation.

Given this introduction to occupations, the next question was
what are new and emerging occupations? The term "new and emerging
occupation" was not officially defined when it first appeared in the
legislation. And it was not seriously considered until a research
was contracted by the United States Office of Education to the
Contract Research Cooperation to study a method of identifying so-
called new and emerging occupations for the purpose of curriculum
development at the national level. The term new and emerging occu-
pation was operationally defined as follows in order to fit the
purpose and conditions of the study:

The new and emerging occupation is one which has come
into existence in the past ten years in skilled or
technical areas, for which there is established demand,
a basis for projecting growth, and a shortage of
trained labor, and for which no public vocational
education is available (Meleen et al, 1976, p. 4).

From the definition specified, there were four conditions:

(1) emerged in the last decade, (2) definite demand, (3) no curriculum
existed, and (4) pertinent to vocational education that operationalized
the term. It may be quite appropriate for a study of this nature to
have defined the term in such a stringent fashion in order that the
identification process can be focused.

Nelson (1979) recently defined it in two separate terms:

New occupation: A job for which the work context and
tasks are substantially new and have been clearly
established during the past two to three years. The
level of demand or potential demand for workers and
sophistication of the tasks warrant vocational training. These jobs may be new at the school district, state, regional, or national level.

Emerging occupation: A series of new job activities in the process of evolving into a job. (p. 1)

His definition of new occupation still followed the restrictive criteria such as time frame of development, level of demand, and pertinency to vocational education training. The definition of emerging occupation was, on the contrary, too broad to have a concrete explanation of what emerging occupation is all about.

New occupations are the new jobs performed due to new functional needs of production and service. These new functional needs may stem from a variety of sources. Some examples may be a new machine design replacing old operation, a new service creating new functions, a new production process requiring new operations, or a new legislation mandating certain new tasks being performed. These new functions, aggregating with some basic or related functions, form a new job entity. Combining the new job entities that have similar or interrelated functions may develop a new occupational title which is distinctively different from the established ones.

Emerging occupations are those lying along the circle with new occupations on the one side and established occupations on the other. They have certain new functional needs but are subsumed under the old job entities. On the one hand, they may become developed as new occupations, provided that the new functions subsumed become increasingly important and are performed at increasing frequencies. They may not be separately developed as new occupations if the new
functions performed are only a small part of the tasks required for
a job. It is conceivably difficult to identify these new and
emerging occupations and even more so to predict whether they will
become new occupations at a certain time in the future.

2.2 ECONOMIC FORCES FOR DEVELOPMENT OF NEW AND EMERGING OCCUPATIONS

There were few empirical studies in the literature that
provide causal explanation. However, there are theories and
postulates in the literature developed under somewhat different
frameworks that do shed some light on the process. Two major
economic forces, one technological change and the other labor market
process, may have significant impact on the development of occu­
pations. (Doeringer and Piore, 1971; Friedrick, 1970; Moore, 1966;
Rothwell and Zegveld, 1979). These factors are treated in more
detail in the following sections.

2.2.1 Technological Change

Technological change is the most frequently mentioned cause
for change in occupational structure (Brozen, 1963; Bureau of Labor
Rothwell and Zegveld, 1979; Taylor, 1965). From the theory of
production, technological change is conceived as a shift in the
production function. The function indicates, for a given level of
technology, the maximum output that can be obtained from a combination
of input, such as capital and labor. The shift in function, due to
technological change, may be labor saving, capital saving, or neutral.
In the literature, the major focus of studies dealing with the
effects of technological change on occupational structure, is on job
displacement effects of automation and its changes in job contents or
skill requirements. We are interested in the effect of technological
change on skill requirements or changes in job contents because those
changes would lead to development of new and emerging occupations.

Technological change will lead to changes in output of the
economy. As the output of the economy changes, the employment in
industry will be shifting, which brings correspondingly shifted mix
of occupation within various industries. The new and emerging occu­
pations will be developed as the "ripple effect" of technological
change on shifting mix of occupations.

The effects outlined have been evidenced by the changes in
occupational structure in the developed economies in the last several
decades. Rapid decline of employment in the primary sector (agri­
culture) is complemented with rapid increase of employment in the
secondary sector (manufacturing) and recently shifting to the
tertiary sector (service).

At the firm level, the impact of technological change on occu­
pational structure follows a similar sequence. Starting from a
possible new major technical innovation, the firm decides to increase
the investment on producing the new products derived from innovation.
This may create a new industry which generates new employment oppor­
tunities. However, the job generation process as suggested by
Doeringer and Piore (1971), may take place when the industry searches
for changes in productive techniques. The industry designs and constructs equipment associated with the new productive technique. Then, the industry designs the jobs required to man the equipment. These newly designed jobs may or may not be substitutions for the old ones depending on the skills required. If a new spectrum of functions is to be performed, it may well be a new job with a new occupational title.

Both the shift in employment structure at the macro-level and job-generating function at the micro-level of the economy are direct effects of technological change. Such shifts not only have effects on the number of jobs but also on the types of jobs that are derived.

2.2.2 Labor Market Process

The literature also suggested that labor market process is an economic force for development of new and emerging occupations (Bureau of Labor Statistics, 1974, 1976; Knox, 1979; Rothwell and Zegveld, 1979). Here labor market process means that, through the interactions of demand and supply of existing occupations, new and emerging occupations may be developed as substitutions for the existing ones. This can be explained in light of the traditional framework of labor demand and supply.

For a given level of supply of existing occupations, higher demand will induce higher wage. This higher wage will eventually force employers to make a decision as to substitute capital investment for higher labor cost on the one hand and to substitute less expensive labor for those increasingly expensive ones on the other.
This decision will subsequently lead to new skill requirements on the part of substituting capital for labor. New skill requirements will eventually lead to development of new and emerging occupations. The development of numerical control machinists is an example in this case. Even for the less expensive labor substituting for increasingly expensive ones, new or emerging occupations with lesser skill requirements will possibly become developed. For example, carpenter I and II perform partially the job functions originally performed by a master carpenter. However, the effect will be reversed if the demand for existing occupations starts to decline, given the level of supply remains constant.

For a given level of demand for existing occupations, higher supply will result in lower wage. This lower wage will definitely not provide employers with the incentive to invest in capital. In addition, a lower wage actually causes the firm to substitute labor for capital. Thus, few new skills will be required. Consequently, few new and emerging occupations will be developed. On the contrary, declining supply of existing occupations will have opposite effect in terms of leading to increased development of new and emerging occupations.

It was noted that in a static framework, both the demand and supply of existing occupations have direct effects on the development of new and emerging occupations. Their effects depend on the assumption of a constant level of demand or supply. This may work in the short run. However, the assumption has to be removed in the long run. When there are changing points of equilibrium between
demand and supply, due to changing quantity or price, the behavior of development of new and emerging occupations mentioned earlier becomes less predictive.

Under the interactions of demand and supply, additional information is needed in order to figure out the appearance of new and emerging occupations. This additional information includes the labor demand and supply functions for existing occupations, the elasticity of substitution between existing and new occupations, the demand elasticity of existing occupations, and the elasticity of substitution between capital and labor in existing occupations. Since few of them are available, alternative measures capturing some dimensions of demand and supply effects will be needed. This will be treated later in the measurement of the variables.

The discussion of substitution between existing and new occupations has so far been on the demand side, which is exerted by the principles of production. Tinbergen (1974) suggested to separate demand and supply sides when dealing with the substitution between various types of labor. Here supply side refers to the willingness of individuals to change from one occupation to the others. This willingness is determined not only by the income, but also by the level of satisfaction with the job. The level of satisfaction will in turn depend on the discrepancy between the level of education required by the occupation and the level of education actually possessed by the workers. A somewhat similar postulate can be raised that the educational level of existing occupations has ramification on whether a worker is willing to choose to work in a new
occupation, which is substitutable for the existing ones. In other words, an individual with a higher level of education in existing occupations will tend to be more flexible to change jobs, maybe move to a new occupation.

2.3 BUILDING A MODEL

Given the theoretical postulates, this section deals with the identification of the variables for the analytic model, the hypotheses postulating the interrelationships among the variables identified, and a proposed conceptual model for the development of new and emerging occupations.

2.3.1 Variables and the Model

It was suggested that technological change and labor market process in terms of its demand and supply effects were the major forces. Four variables were identified relevant to technological change. They were research and development in government, research and development in private companies, sectoral growth, and occupational structure change. Also identified for labor market process were demand and supply effects. These variables were described in the following sections.

It was noted that the impact of technological change on the development of new and emerging occupations was not a direct one. Evidence has shown that sectoral growth has effects on technological change and vice versa (Kendrick, 1973). Sectoral growth will result in increase in output to the economic sector. The profit from
additional output may be spent proportionally on capital improvement, a form of technological change. This technological change will lead to increases in production efficiency, which subsequently leads to increases in income. This increase in income will lead to increase in demand, which eventually leads to sectoral growth.

Evidence has also shown that sectoral growth has effects on occupational structure change (Rothwell and Zegveld, 1979). Here occupational structure change refers to changes in the proportion of occupational employment in an industry and in the economy. The effects of sectoral growth on economic sectors are twofold. The first-order effect is the increased level of demand and in turn increasing the level of employment. However, this is not the only possibility. The level of employment may not be up even though the level of demand is increased because the increased demand can stimulate capital investment in machinery, which may, on the contrary, reduce the number of workers needed. But this reduction in work force must accompany more likely the advanced skills mix of the work force. And, consequently, it may create new types of occupations in order to man the new machinery. The second-order effect is the relocation of economic sectors as evidenced by the increasing share of service sectors and continuous decline of agricultural and manufacturing sectors in the developed economies (Rothwell and Zegveld, 1979). This relocation effect, however, can be subsumed under increased level of demand since demand increased in one sector (or industry) may follow concurrently a decreased (or saturated) demand in another. Thus, the occupational structure will be changed in terms
of its quantity and quality. This change is consistent with the changing levels of employment, which depend on the direction of demand. In other words, both the number and the type of occupations will be changing. However, their respective effects are difficult, if not impossible, to quantify. No evidence has been found that sectoral growth has a direct effect on the development of new and emerging occupations. Its effect may have to go through the structural change in occupations. Therefore, the causal link between sectoral growth and development of new and emerging occupations can only be established via the variable of occupational structural change.

The effect of technological change on occupational structure change, which leads to development of new and emerging occupations, was mentioned. However, this only reflected the possible impact on the type of new occupations being distributed in the economic sector and partially the impact on the quantity of new occupations. The major impact of occupational structure change on the quantity change of new and emerging occupations is propagated through the demand effects of substitutable occupations. This is where the labor market process interacts with the force of technological change.

It seemed that the variable of technological change has so far been treated as an exogenous variable. In fact, technological change is rather endogenous to the process. There are a number of factors influencing the rate of technological change. Some of them are the expected profitability of investment in research and development (hereinafter referred to as R & D), cost reduction of a particular
product, increased demand for a particular product, shortage or price rise of a production factor, market competition, the attitudes of workers, management, and the public toward technological change, process of science and technology resulted from R & D, and government support on R & D (Mansfield, 1968). A recent study was done by Kendrick and Grossman (1980) using total factor productivity growth (an index representing technological change) within twenty manufacturing industries to regress against a number of variables. The variables were capacity utilization, education, research and development expenditures, hours worked per week, degree of concentration within the industry, degree of unionization, percent of man-days idled due to work stoppages, composition of work force by sex, type of worker, layoff and quit rates, and shifts of employment among industries at two time periods of 1948-66 and 1966-76. The purpose was to find out the causal factors for the growth of total factor productivity. The result indicated that only research and development expenditure had a significant and stable effect on productivity over the two periods.

It is recognized that research and development has direct contributions to raising factor productivity which is a reflection of technological change. The direct contribution is direct increase in productivity of privately or government-financed industries or organizations conducting the R & D. The indirect contribution is indirect increase in productivity of industries purchasing capital and intermediate inputs (embodied technology) from the industries or
government organizations conducting the R & D. Higher estimates were found in indirect effects of privately financed R & D than those in direct ones, especially in manufacturing. No significant effect, either direct or indirect, was found for government-financed R & D (Terleckyj, 1974). Even after taking human capital (measured as investment intensity ratio based on growth of real wages per man-hour worked), another form of input, in the estimation equations, similar results were obtained (Terleckyj, 1980).

One result, which is nonsignificant contributions, both direct and indirect, of government-financed R & D to productivity growth, deserves some discussion. It is well recognized that the benefits derived from R & D tend to be realized only over time and the risks are quite high in terms of the return of investment. Hence in the modern nations, government generally takes the initiative and risk to fund R & D projects which may not have immediate impact on the factor market. Globerman (1980) pointed out that federally financed R & D is primarily directed toward improvement in product quality as opposed to cost reduction. In other words, federally funded R & D projects are aimed at facilitating changes in production conditions. Therefore, it may take a longer period for any effects on productivity growth to be significant. For this reason, government investment in R & D was included as an exogenous variable in the analysis.

Little evidence has been found in terms of the direct effect of investment in R & D on change in occupational structure. However, it is intuitively clear that the effects of investment in R & D on the
development of new and emerging occupations must be propagated through several stages of development; for instance, sectoral growth and structural change in occupations.

In addition to R & D investment, there are a host of exogenous variables having different impacts on technological change as suggested by Kendrick and Grossman (1980) and Terleckyj (1980) in their productivity studies. However, their respective significant contributions are not conclusive. Besides, little evidence has been documented about the effects on the other endogenous variables considered in the study. Thus, only R & D investment variables which have constant effect on technological change were included.

In summary, five variables were identified under the force of technological change. They are R & D in government and private companies, technological change, sectoral growth, and occupational structure change. Two variables were identified under the force of labor market process. They are occupational demand effects and occupational supply effects. Their causal relationships were described earlier. The hypotheses postulating the causal relationships will be treated in the next subsection.

A conceptual model representing the causal relationships among exogenous and endogenous variables were established. The model is depicted in Figure 2.

The arrows in Figure 2 represent the directions of causal inference being proposed in the hypotheses. The magnitude of these "arrows" would be the parameters to be estimated. Their significant contributions would be tested later in the study.
Figure 2: A Conceptual Model for Development of T & I New and Emerging Occupations

RDG :  R & D in Government
RDP :  R & D in Private Companies
TC :  Technological Change
SG :  Sectoral Growth
OSC :  Occupational Structure Change
ODE :  Occupational Demand Effect
OSE :  Occupational Supply Effect
NEO :  New and Emerging Occupations
2.3.2 The Hypotheses

Based upon a summary of the discussions in the earlier subsections, several hypotheses were formulated concerning the causal relationships among the variables identified. It is important for the reader to recall that only T & I new and emerging occupations in manufacturing were included in the study. Nine hypotheses were made in the study.

1. The demand effect of existing occupations has a significant direct effect on the development of T & I new and emerging occupations in manufacturing.

2. The supply effect of existing occupations has significant direct effect on the development of T & I new and emerging occupations in manufacturing.

3. The occupational structure change has significant direct effect on the development of T & I new and emerging occupations in manufacturing.

4. The occupational structure change has significant indirect effect on the development of T & I new and emerging occupations in manufacturing through the variables of occupational demand effect and occupational supply effect.

5. There exists a significant reciprocal causal relationship between the occupational demand effect and occupational supply effect variables in manufacturing.

6. The sectoral growth variable has significant direct effect on the occupational structure change variable in manufacturing.
7. There exists a significant reciprocal causal relationship between technological change and sectoral growth variables in manufacturing.

8. The technological change variable has a significant direct effect on occupational structure change in manufacturing.

9. Research and development in both the government and private companies conducted in manufacturing have significant direct effects on technological change.

2.4 FROM CONCEPTUAL MODEL TO MULTIPLE INDICATORS MODEL

One of the purposes of this study was to verify empirically the model and to make refinements if necessary. Here empirical verification takes the process of testing the parameters estimated. However, in order for the parameters to be estimated, the variables must be measurable. And also, the data must be available. While looking at the model in Figure 2, the researcher noted that most of the variables were theoretical constructs or latent variables. For instance, the variables of technological change, occupational structure change, and demand and supply effects are not directly measurable. Thus, the problems of measuring those variables needed to be resolved before any empirical verification could be performed. This was exactly the mission of this section to translate from a conceptual model to an empirically verifiable one. This section dealt with the measurement of each variable and the selection of indicators. Also presented in this section was a multiple indicators model that is empirically verifiable.
Here, a multiple indicators model was implied as the target model to be verified empirically. However, the details for the rationale behind this selection would not be presented until the later chapter of methodological considerations.

2.4.1 Measurement of the Variables

One obvious concern was the measurement of the dependent variable of new and emerging occupations. It was defined as the change in distribution of T & I new and emerging occupations in manufacturing. The new occupational titles contained in the fourth edition of the *Dictionary of Occupational Titles* (hereinafter referred to as *DOT*) published in 1977 as compared to those in the third edition published in 1965 were used to represent new and emerging occupations. They were selected because the coverage of occupational titles was based on detailed occupational analysis done by the Department of Labor through its eleven occupational analysis field centers in the United States. In order to translate those new occupational titles into equivalent educational program areas, the crosswalk between the fourth edition *DOT* code and vocational education program code available from the Bureau of Labor Statistics (1980) was used.

In general, there were two units of measurement for the employment variable. The first was to use the (absolute) employment figure of each new and emerging occupation as the unit of measurement. However, it had two problems. First, the employment figures of different new and emerging occupations were not homogenous in
terms of the range of value. One occupation may have had a very small number of employees while the other may have employed quite a large number. Thus, the analysis was very sensitive to those extreme employment figures. Second, there was no base of comparison among those T & I new and emerging occupations in terms of their relative contributions to the employment in manufacturing. Since the aggregate or the macro-picture rather than the individual or micro-structure of the developmental behavior was of interest, the second problem became salient.

The second way to quantify the variable was to use percentage distribution of the employment of each T & I new and emerging occupation to the employment of the manufacturing as the unit of measurement. This measure overcame the difficulties mentioned above. In addition, this pattern of percentage distribution could show the cross-sectional structure of the development of T & I new and emerging occupations in manufacturing at a certain point in time. Therefore, the second way of quantification appeared to be more appropriate to this study. Incidentally, this variable is not a latent but a measured one.

The second variable was the demand effect. This variable is a latent variable. Thus, indicators were needed. As indicated earlier, in order to quantify the demand effect of those existing occupations on the new and emerging ones, two dimensions of this effect, price and quantity, should be considered. For occupational demand effect, the average annual change rates of demand of occupations being
substitutable by those new and emerging ones were used to reflect the quantity dimension. Here demand means the net opening of the occupation. An implicit assumption of unity elasticity of substitution between existing and new occupations was made.

Next, the prices for substitutable occupations was considered. Here price refers to the wage level. It is conceivable that the higher the demand of those existing occupations being substituted by new and emerging ones is, the higher the wage level will become, given that the supply remains at the same level. If the wage level for existing occupations remains high due to high demand, it will affect the appearance of those new and emerging occupations. Thus, the average annual change rates of the wage levels of existing substitutable occupations can reflect one dimension of the effect of this variable and were used as the second indicator.

The third variable was the supply effect. This is an endogenous latent variable. It has been indicated by Marshall, King, and Briggs (1980) that at least two dimensions should be considered in the effect of labor supply. One is quantitative dimension which deals with the amount of labor supplied while the other is qualitative dimension which deals with skills composition and workers' characteristics of the labor market. Thus, two indicators were used to reflect these two dimensions of occupational supply effect. The first one was the average annual change rates of the supply of occupations being substitutable by the new and emerging ones. Here supply means the number of individuals in a particular occupation available to the labor force at a particular
point in time. This reflects change in quantity of labor supply. The second one was the average annual change rates of educational attainment of those existing occupations. This reflects change in quality of labor supply.

The fourth endogenous variable was the variable of occupational structure change. This is a latent variable. There are two dimensions of structural change in occupations. One stems from the interaction effect of changing demand and changing skill requirements. The other stems from only the simple effect of changing demand. Therefore, two indicators were proposed. The first was the occupational structure shift index which is defined as the difference of the Industry-Occupation Matrix Ratios of an occupation between two time points. The Industry-Occupation Matrix ratio is the ratio of employment in an occupation to the employment in an industry to which this occupation belongs. This shift index represents the pattern of structural change of an occupation in a particular industry as the result of either changing pattern of demand or changing skills requirement in that industry.

The second indicator is called occupational employment shift index. This shift index is defined as the difference of occupational employment ratios of an occupation between two time points. The occupational employment ratio is the ratio of the employment in an occupation in its belonging industry to the employment in that occupation in the entire economy. This shift index implies the pattern of employment change of an occupation of a particular industry in the entire economy as the result of changing pattern of demand.
There is a subtle difference between these two indices. The structure shift index is mainly concerned with the change of occupational distribution within an industry while the employment shift index is concerned with the amount of occupational employment change in an industry. The former roughly represents both the effects of changing demand and changing skill mix while the latter represents only the effect of changing demand. The units of measurement for those two indicators were the average annual change rates of these indices with respect to the occupations being substitutable by the new and emerging ones. The mathematical representations of these two measures will be found in the next chapter of data considerations.

The fifth endogenous variable was the variable of sectoral growth. Sectoral growth can be indicated in two ways. The first one is the average annual growth rate of individual manufacturing industry related to the new and emerging occupations identified. This measure corresponds to the effects (both the first and second order) of increased level of demand due to sectoral growth as indicated earlier. The second indicator is the average annual change rate of capital/labor ratio in industries related to the new and emerging occupations of concern. This corresponds to the other path that increased demand may trigger more investment in capital. This may result in changes in skill mix, which in turn results in changes in occupational structure on the one hand and technological change on the other.
The sixth endogenous variable, technological change, deserved rather extensive discussion. This one is also a latent variable. Measuring technological change has long been a major controversy in the economics of production. The major indicator used to represent technological change has been the growth in productivity. However, the measurement of growth in productivity itself sparked rather lengthy theoretical and practical discussions. Nadiri (1970) did a thorough survey on this issue. Productivity is in general to denote the relationship between outputs and the associated inputs used in the production process. But its measure may vary depending on the definition of output and input, the weighting patterns used to combine different units of outputs and inputs, and the manner in which outputs are related to inputs.

One of the simplest ways in technological change measures is to use labor productivity \( (P_L) \), which is output per man-hour of labor \( (Q/L) \). It is likely that a rapid rate of technological change is to result, all other things being equal, in a high rate of growth of labor productivity. Albeit that technological change is the major determinant of labor productivity, there are other important factors influencing the growth rate of labor productivity. For example, the extent to which capital is substituted for labor in response to changes in relative input prices will increase labor productivity. Also, the economics of scale, the rate of diffusion of best practice, and the nature of technological change (capital saving vs. labor saving) are major considerations in labor productivity improvement.
Despite it inadequacy in measuring technological change, labor productivity has long been the official index used by the Bureau of Labor Statistics in productivity release. It reflects the joint effects of new technology, capital investment, level of output, capacity utilization, energy use, managerial skills, and skills and efforts of the work force. Although the contributions of the effects are considered very implicitly, Clague (1966) brought out his objection to using productivity as a measure of technological change for two reasons: (1) technological change alters product content over time, hence, a productivity measure, based on a particular product content, is valid only for a short period of time, and (2) productivity is usually underestimated because the output is underestimated due to steadily falling prices.

Recognizing the inadequacy of official index in productivity measure, the researcher concluded it to include explicitly the capital as a contributing factor input in the measure. This is so-called Total Factor Productivity (hereinafter referred to as TFP). Two TFP measures often used are Kendrick's arithmetic measure and Solow's geometric measure. The former is based on distributional theory where the production function is implicit while the latter is explicitly deduced from a Cobb-Douglas production function. Their respective measures are simply stated as follows:

\[ K_{tp} = \frac{Q}{a_L + (1-a)K} \] (2.1)
where Q : aggregate level of output adjusted for some base year
L : aggregate level of labor input adjusted for some base year
K : aggregate level of capital input adjusted for some base year
a : labor's share of the value of output in the base year estimated from income account
1-a : capital's share of the value of output in the base year estimated from income account

Solow's measure
\[ S_{\text{tfp}} = \frac{Q}{L^bK^c} \]
where Q, L, K are the same as the above
b : elasticity of output to labor
c : elasticity of output to capital

Of course, both measures require the assumption of disembodied technological change (no quality change in labor and capital input along the time) and constant return of scale. In addition, the geometric measure requires the assumptions of unity of substitution between labor and capital and a homogenous production function while arithmetic measure requires no diminishing marginal productivity of individual factors (Kendrick, 1973).

In addition to the restrictive assumptions, it is apparent that TFP is not measuring the "pure" technological change since only tangible factors, despite their dominant importance, are included. Other intangible factors such as quality change of labor and capital due to education and health improvement and better product mix and resources allocation, are treated as unknown in the "residual" of
this measure. A study done by Kendrick and Grossman (1980) in an effort to differentiate their respective contributions to productivity growth yielded no conclusive results.

Since there was no decisive theoretical reason to choose one measure over the other, the selection came down to empirical parsimony. Kendrick (1973) indicated that these two measures are interchangeable if the component variables and the underlying shape of the relationship are consistently specified. He also indicated that the elasticity of substitution between labor and capital in the United States is significantly below 1.0, which is contrary to the assumption in Cobb-Douglas production function from which the geometric measure is derived on the one hand and the arithmetic measure can reflect the changing factor shares of income on the other. Besides, the data for factor shares estimated from national income account are readily available. Therefore, for the purpose of this study, the average annual change rate of this arithmetic index during the time period of concern (1965-1977) was used as the measure for productivity growth, which is an indicator of technological change.

Another possible measure used as a crude index of technological change is patent intensity ratio. This ratio is defined as the number of patents granted to an individual manufacturing industry compared to the total number of patents granted in the entire manufacturing industry (Mansfield, 1968; Schmookler, 1972). This indicator is not independent of economic variables. Evidence
suggested that a high correlation exists between a high patent rate on capital goods inventions and a high previous level investment or value-added. This high correlation can be explained both by demand and supply factors (Schmookler, 1966). However, the importance and the cost of the patents are treated equally in these statistics. It is extremely difficult, if not impossible, to weigh different patents in terms of their impact of economic contribution less their costs. Therefore, the average annual change rates of this patent intensity ratio in different manufacturing industries were chosen as the second indicator for technological change.

The seventh and eighth variables, R & D investments in government and private companies, were exogenous ones. They are not latent variables and can be measured directly. Terleckyj (1974, 1980) employed R & D investment intensity ratio, which is defined as the ratio of gross R & D investment in an industry to the net sale of that industry, as the measurement in his study. The similar measures were used to indicate the contribution of R & D investment to technological change. For government R & D investment, the average annual change rates of government-financed R & D intensity ratios conducted in manufacturing were used. Similarly, the average annual change rates of privately financed R & D intensity ratios conducted in manufacturing were used for private companies R & D investment.

2.4.2 The Multiple Indicators Model

From a summary of the materials presented in the subsections of conceptual model and measurement of variables, two observations
surfaced. First, there exist reciprocal relationships among variables, thus, simple multiple regression does not capture the postulated relationships. Second, some of the variables are not directly measured, thus, indicators are necessary. These conditions led to the formulation of a multiple indicators model. The variables, indicators, and their mnemonics are presented in Table 1. The model bearing the mnemonics is depicted in Figure 3. It should be noted that Figure 3 was an expanded version of Figure 2 with indicators representing their respective latent variables.

2.5 SUMMARY

This chapter included four sections. The first section described the meanings of new and emerging occupations. The second section described the economic forces underlying the development of new and emerging occupations. The third section built a conceptual model based on the theories and postulates described in the second section. The last section expanded from a conceptual model to a multiple indicators model which could be verified empirically.
<table>
<thead>
<tr>
<th>Exogenous/Endogenous Variables</th>
<th>Indicator</th>
<th>Mnemonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government R &amp; D investment (ex)</td>
<td>Average annual change rate of government R &amp; D expenditure intensity ratio</td>
<td>CRDG</td>
</tr>
<tr>
<td>Private company R &amp; D investment (ex)</td>
<td>Average annual change rate of private company R &amp; D expenditure intensity ratio</td>
<td>CRDP</td>
</tr>
<tr>
<td>Technological change (en)</td>
<td>1. Average annual change rate of Total Factor Productivity</td>
<td>CTFP</td>
</tr>
<tr>
<td></td>
<td>2. Average annual change rate of patent intensity ratio</td>
<td>CPIR</td>
</tr>
<tr>
<td>Sectoral growth (en)</td>
<td>1. Average annual change rate of industry output</td>
<td>CINO</td>
</tr>
<tr>
<td></td>
<td>2. Average annual change rate of capital/labor ratio</td>
<td>CCLR</td>
</tr>
<tr>
<td>Change in occupational structure (en)</td>
<td>1. Average annual change rate of occupational structure change</td>
<td>COSC</td>
</tr>
<tr>
<td></td>
<td>2. Average annual change rate of occupational employment change</td>
<td>COEC</td>
</tr>
<tr>
<td>Demand effects (en)</td>
<td>1. Average annual change rate of wage levels of existing occupations</td>
<td>CWLE</td>
</tr>
<tr>
<td></td>
<td>2. Average annual change rate of demand of existing occupations</td>
<td>CODE</td>
</tr>
<tr>
<td>Supply effects (en)</td>
<td>1. Average annual change rate of supply of existing occupations</td>
<td>COSE</td>
</tr>
<tr>
<td></td>
<td>2. Average annual change rate of educational levels of existing occupations</td>
<td>CELS</td>
</tr>
<tr>
<td>Change of new and emerging occupations (en)</td>
<td>Change in distribution of T &amp; I new and emerging occupations in manufacturing</td>
<td>CNEO</td>
</tr>
</tbody>
</table>
Figure 3: A Multiple Indicators Model for Development of T & I New and Emerging Occupations
Chapter III

DATA CONSIDERATIONS

The theories, measurement, and models were discussed in the previous chapter. It seemed that empirical verification should be the next concern. Empirical verification of the causal model was one of the purposes in the study. This verification must be done with data. Results of model testing and sources for model modification relied heavily on the quality of the data collected. Thus, special attention was given to data in this chapter.

In traditional educational research, either experimental or non-experimental, data are collected mainly via survey or testing. The objects from which the data are obtained are human beings. Individual behavior, perception, reaction, and interaction on certain inquiries or events are of interest. Social science research including educational research should be viewed from a broader perspective in which inquiries on both individual human behavior and collective human behavior are included. Here collective human behavior means a behavior reflecting or a phenomenon resulting from collective human actions. The phenomenon rather than the individual action is of interest. For example, the economic behavior resulting from expected inflations, the social mobility resulting from expected inflations, the social mobility resulting from socioeconomic status
of the family, and the voting behavior reflecting different political climates, are a few. For collective behavior, information has generally been gathered from individuals and been aggregated in certain forms by either government or other data producers. These data may be collected for a particular purpose and may be transformed in some fashion before any other use. Nonetheless, they are collected by professionals and their quality should be maintained at least at an acceptable level.

The study at hand fell into the latter category since the development of new and emerging occupations is regarded as a collective behavior. Thus, the data were collected mainly from published sources pertinent to the variables of concern. The following sections deal with population and sample, selection of substitutable occupations, sources for the data, and data transformations or adjustments wherever appropriate.

3.1 POPULATION AND SAMPLE

First, the population of the study was examined. Although the term new and emerging occupation was not defined precisely in the study, a good practical approximation to that can be found from the additions of occupational titles in the fourth edition of the Dictionary of Occupational Titles (DOT) in 1977 as compared to the third edition in 1965. More than 2,100 new occupational titles in the fourth edition were not listed in the third one (Employment and Training Administration, 1977). Presumably, these new titles represent
new and emerging occupations developed during the time period between 1965 and 1977. Since the compilation of the DOT represents massive efforts done by those occupational analysis field centers in the United States, the information should be most appropriate, if not the best available, to this study.

After the identification of the population, an accessible frame should be in place for sampling. Two criteria were used for determining the sampling frame. First, the occupations in the frame must be appropriate for education and training in T & I areas. The appropriateness was judged based on the Special Vocational Preparation (SVP) ranging from coding numbers 3 to 8 (three months to four years needed for satisfactory performance). Second, there should be occupational employment statistics available for those occupations in the frame.

The population could be obtained from a recent publication dealing with conversion in codes and titles changes between the third and fourth editions of the DOT by the Employment and Training Administration (1979), in which the codes and titles for new occupations are listed. This narrowed the work down to dividing from the list the new occupational titles appropriate for T & I training on the one hand and to examining their occupational employment statistics on the other. For the former criterion, the problem came down to assessing equivalent vocational preparation for those new occupations. This did not impose any problem for the third edition DOT since crosswalk between the third edition DOT and vocational education programs codes was developed (Office of Education and
Manpower Administration, 1969; California Manpower Management Information System, 1976). However, for the fourth edition DOT, complete crosswalk was yet to be published. The National Occupational Information Coordinating Committee (1979) has published partial crosswalk between the third and fourth editions of DOT on the one hand and between DOT and vocational education program codes on the other. This crosswalk is far from complete. For the second criterion, the problem became even more difficult. DOT is by far the most detailed occupational classification for the United States' economy (more than 20,000 occupations). But its corresponding employment statistics are too expensive to collect and may be too detailed to be useful for most data users. Traditionally, the employment statistics using Census occupational classification scheme are the most comprehensive statistics available. However, this scheme only has 440 occupational groups. It suffers, on the one hand, from the criticisms of being too aggregated to tell the insights of occupational employment of the economy and from using socioeconomic attributes rather than job function as the classification on the other (Scoville, 1969).

An Occupation Employment Statistics (OES) program was launched by the BLS in 1971 to collect occupational data at the classification level between the DOT and the Census. This OES program is a federal-state cooperative and has three phases of operation: survey, matrix, and projection (Bureau of Labor Statistics, 1976). The OES survey contains about 2,000 occupations based on DOT and Census.
classifications. The national survey was conducted over a three-year cycle for manufacturing in one year and for non-manufacturing industries in the other two years. These employment data are by far the most detailed ones available in a finer occupational classification system.

Judging from the two criteria mentioned, the accessible frame could only be obtained by matching those new occupations, on the one hand, with the Special Vocational Preparation in T & I areas and with the OES Occupational classification on the other. The latter needed a more comprehensive crosswalk among different classification systems. Fortunately the BLS has recently completed a crosswalk tape with different occupational systems so that the matches could be performed (Bureau of Labor Statistics, 1980).

A total number of 419 new T & I occupations qualified for Special Vocational Preparation and employment statistics criteria. However, two problems emerged which required further refinement of the frame. First, there are new and emerging occupations that have generic OES titles and codes (i.e., 55B29 for all supervisors of non-workers, 59001 for all other skilled workers, 59002 for all semiskilled workers, and so on); which made the distinction between new and established occupations impossible. Second, a great number of new occupations in the frame had employment less than 0.5% of the employment in the industry to which the occupations belong. And no employment statistics were published under this situation. Excluding these "contaminations" in the frame, a much smaller number of T & I
new and emerging occupations (98) were left for analysis. The final inclusions in the frame are listed in the Appendix. It was decided to use all of them in the study since the number was too small to allow for meaningful sampling.

3.2 SELECTION OF SUBSTITUTABLE OCCUPATIONS

The term, "substitutable occupations," refers to the subset of existing occupations being substitutable by new and emerging ones. Apparently some criteria should be used to determine their selections. Here the criteria employed were highly problematic to the study due to the availability of data.

Four factors representing different dimensions of an occupation were considered in the selection of substitutable occupations out of the vast number of occupations in the labor market. First, the substitutable occupations should have the same first three-digit DOT code as the new and emerging one. This three-digit code represents a finer category of occupational group in which similar job functions are performed. Second, the middle three-digit DOT code of the substitutable occupations should be equal to or not greater than those in the new and emerging ones. On the one hand, the middle three-digit code represents, respectively, the relationship of the worker to data, to people, and to things in the performance of the job and with 0 being the most complex level of relationship and 8 the most simple one. On the other hand, the judgment that a job
performance relationship with higher scale should be substitutable for the lower ones, in the same job function category, can be made. Third, the substitutable occupations must have at least the same level of Special Vocational Preparation as the new and emerging ones. It was felt that in order for an occupation to be perfectly substitutable, no less than the amount of special vocational training needed to perform the job functions is a sufficient condition. Lastly, and perhaps the most critical consideration, was that there must be employment statistics available for those substitutable occupations so that empirical verification of the model in the later stage can be performed. This last constraint limited the considerations of substitutable occupations to those being surveyed in the OES program. Nevertheless, based on the criteria listed above, substitutable occupations for each new and emerging occupation included in the frame could still be obtained.

3.3 DATA FOR ENDOGENOUS VARIABLES

This section is a description of the sources and data transformations for the indicators of those endogenous variables mentioned earlier. It is organized into subsections. In each subsection, the data considerations for one or more endogenous variables were elaborated.

3.3.1 Data for New and Emerging Occupations

Collection of employment statistics was a crucial step for the T & I new and emerging occupations included in the frame. It was
indicated earlier that the OES is by far the most elaborate effort
done in the United States in collecting detailed national occupational
employment statistics. Its survey results would certainly serve the
purpose of this analysis. The crosswalk mentioned earlier was used
to cross-reference the occupational codes between DOT and OES.
Since T & I occupations fall almost exclusively in manufacturing, the
occupational employment statistics based on three-digit SIC manu-
facturing industries codes were used to obtain the measures for new

The measures for the endogenous variable of T & I new and
emerging occupations are shown in the following matrix:

<table>
<thead>
<tr>
<th>SIC</th>
<th>$Y_1$</th>
<th>...</th>
<th>$Y_j$</th>
<th>...</th>
<th>$Y_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>$r_{11} w_1$</td>
<td>...</td>
<td>$r_{1j} w_1$</td>
<td>...</td>
<td>$r_{1n} w_1$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_i$</td>
<td>$r_{il} w_i$</td>
<td>...</td>
<td>$r_{ij} w_i$</td>
<td>...</td>
<td>$r_{in} w_i$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_m$</td>
<td>$r_{ml} w_m$</td>
<td>...</td>
<td>$r_{mj} w_m$</td>
<td>...</td>
<td>$r_{mn} w_m$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sum_i r_{il} w_i$</td>
<td>...</td>
<td>$\sum_i r_{ij} w_i$</td>
<td>...</td>
<td>$\sum_i r_{in} w_i$</td>
<td></td>
</tr>
</tbody>
</table>
The symbols and their transformations are:

\[ r_{ij} = \frac{E(Y_j)}{E(X_i)} \]  
\[ w_i = \frac{E(X_i)}{\sum E(X_i)} \]  
\[ r_{ij}w_i = \frac{E(X_j)}{\sum E(X_i)} \]  
\[ \frac{\sum r_{ij}w_i}{\sum w_i} = \frac{m}{\sum E(Y_j)} \frac{m}{\sum E(X_i)} \]

where \( r_{ij} \): proportion of new and emerging occupation \( Y_j \)'s employment to industry employment of industry \( X_i \)

\( w_i \): proportion of industry \( X_i \)'s employment to the total industry employment of manufacturing industry \( m \)

For new and emerging occupation \( Y_j \), the measure is equation (3.4), which shows the proportion of new and emerging occupation \( Y_j \)'s employment to the total employment in manufacturing. In the meantime, the changes in distributions of these new and emerging occupations for the period of concern are essentially equal to the measure in (3.4). Here the assumption of zero employment of these new and emerging occupations in the beginning was made.

3.3.2 Data for Occupational Demand Effect

The occupational demand effect of those substitutable occupations on the new and emerging ones was represented by two indicators. One was the average annual change rate of demand of substitutable occupations related to the corresponding new and emerging one. Here demand refers to the net openings of substitutable occupations. The other was the average annual change rate of their
wage level. Since there might be multiple occupations being
substitutable by one new and emerging occupation, the weighted
average of change rates of these substitutable occupations with
the employment as the weighting factor was used. It is shown in
the following formula:

\[ D_j = \frac{\sum_{k} (D_{ki}E_k/\sum_{k} E_k)}{n} \]

(3.5)

where \( D_j \) : weighted average annual change rate of
demand of substitutable occupations for
new and emerging occupation \( j \)

\( D_{ki} \) : average annual change rate of demand of
substitutable occupation \( k \) in industry \( i \)

\( E_k \) : employment of substitutable occupation \( k \)

\( n \) : number of substitutable occupations for
new and emerging occupation \( j \)

It was assumed that the annual change rates of demand of
substitutable occupations remain stable so that a constant average
annual change rate can be calculated for the time period of concern.
Two methods, the least squares trends of the logarithms of the demand
and the heuristic approach of averaging the ratios of the differences
between two consecutive years to the base year, were used to figure
out these change rates. If a decent fit \( (R^2 \geq 0.7) \) was obtained, the
least squares ratios were adopted. If a decent fit was not obtained,
the heuristic ratios were utilized. On the one hand, a poor fit of
the regression line to the data represented a poor selection of
change rate for the data. On the other hand, the heuristic ratio
did have the appeal of conforming to the definition of average
annual change rate used in the indicator. However, the decision of 0.7 being the cutting edge was still highly subjective.

Since the BLS has just published its first round of data collection on OES, it was impossible to generate any time trends out of this program. Therefore, making the best of existing statistics appeared to be the most rational alternative. The source of data for change rates calculation was from the BLS's employment projection program in which the net openings of the occupations within each industry were included. These occupational statistics are in Census code; so crosswalk between OES and Census for those substitutable occupations was needed. It was recognized that the time period covered in the data (1970-1976) was far from complete in providing the full spectrum of occupational employment trends. However, if the assumption of constant change rates was to be maintained, the estimates should be close to the true values.

Next, obtaining the average annual change rates of wage levels for those substitutable occupations was not a simple task. First of all, no wage data were available at the occupational level. The most detailed wage data in the United States are the Employment and Earnings series containing wage data at industry level put out by the BLS (1978). So the assumption that the average annual change rate of wage level of substitutable occupation is close to that of the industry within which the substitutable occupation is distributed had to be made in order to use the data. Of course, the finer the industry level was, the better the approximation would be. Three-digit SIC industry level consistent with the employment statistics
used for new and emerging occupations mentioned earlier was sufficiently detailed under this circumstance. Linear least squares or heuristic approach—whichever was better, as mentioned earlier—was used to calculate the change rates. Similarly, the change rates of multiple substitutable occupations were weighted according to their employment. The formula for calculation was the same as (3.5) only substituting \( W_i \) and \( W_j \) for \( D_{ki} \) and \( D_j \) respectively.

3.3.3 Data for Occupational Supply Effect

It was mentioned earlier that occupational supply effect of substitutable occupations can be indicated by two dimensions. One is concerned with the effect of quantity change being indicated by the average annual change rates of supply for those substitutable occupations. The other represents the effect of quality change, as indicated by the average annual change rates of educational attainment related to those substitutable occupations. The first indicator could be approached from the standpoints of "stock" and "flow" of supply. The supply stock indicates the number of individuals in a particular occupation available to the labor force at a particular point in time. The supply flow indicates the net number of individuals adding to and subtracting from the stock due to new entrants and withdrawals from the labor force and mobility among occupations. The supply flow in previous period results in the present stock (Bureau of Labor Statistics, 1974). Thus, the use of the change rate of supply stock could capture better the quantity dimension of occupational supply effect.
It was recognized that data collected for stock of occupational supply are far from satisfactory. The statistics closest to this need were obtained from the occupational employment series based on Census classification in the Handbook of Labor Statistics, 1978 edition. The series covered only the time period from 1972 to 1977, which may not capture the complete picture of change. However, if the assumption of constant annual rate of change could hold, the statistics could still be useful. The change rates were obtained either by least square or heuristic approach, depending on which one was better. Similar to the demand effect, the weighted average change rates were calculated using formula (3.5) except substituting the change rates on supply of substitutable occupations for $D_j$.

For the quality dimension of occupational supply, the change rates on median year of education can represent the change in educational attainment which may facilitate the development of new and emerging occupations as proposed earlier. However, the availability of those change rates on the median year of education at occupational level is open to question. The Educational Attainment of Workers series, published by the BLS, provide annual data on the median year of education at industry level. In order to use the data, an assumption had to be made that the median educational requirement of a manufacturing industry can represent, to a certain extent, the educational requirement of a substitutable occupation belonging to that industry. It is not an unreasonable assumption, since the substitutable occupations are falling into vocational
training which generally lies in the middle of the educational requirement spectrum in manufacturing. The change rates of the educational attainment were calculated in the same manner as those for the change rates of the wage levels in the occupational demand effect variable.

3.3.4 Data for Occupational Structure Change

For the variable of occupational structure change, the average annual rates of change in occupational structure shift index and in occupational employment shift index with respect to the substitutable occupations were suggested as the indicators. It was noted earlier that the occupational structure shift index is the difference of the Industry-Occupation Matrix ratios (occupational employment of a substitutable occupation to its industry employment) between two time points. The change rate of this index is measured as the proportion of this index to the base year ratio. Since this occupational structure shift index is relatively insensitive to annual change, only periodical data were available. The BLS has put out data on national I-O Matrices for 1960, 1970, 1976, and 1978, and projected matrices for 1975, 1980, 1985, and 1990, from which the occupational shift index can be calculated. It was recognized that these data do not cover exactly the same time period as the one in this study. The limitation of data availability made it necessary to go with whatever appeared compatible. It was decided that only the 1970 and 1976 matrices would be used in calculating the change rates of the index. In addition, the BLS matrix data have occupational classifications very similar but not exactly equivalent to
the Census classification. Therefore, crosswalk among OES occupation, BLS matrix occupation, and Census occupation was used to match the data.

The formula for calculating the average annual change rate of the occupational structure shift index is as follows:

\[ I_j = \frac{1}{m} \sum_{i=1}^{m} \left( \frac{E_i / E_{t2}}{E_i / E_{t1}} \right) \left( I_i - I_i \right) / (I_i)(t_{2}-t_{1}) \]  \hspace{1cm} (3.6)

where \( I_j \) : average annual change rate of occupational structure shift index of substitutable occupations related to new and emerging occupation \( j \)

\( t_1 \) : occupational matrix ratio of substitutable occupation \( i \) at time \( t \)

\( I_i \) : occupation employment of substitutable occupation \( i \)

\( E_i \) : occupation employment of substitutable occupation \( i \)

\( m \) : number of substitutable occupations for new and emerging occupation \( j \)

The second indicator of the average annual change rate of occupational employment shift index can be obtained in a similar way. As indicated earlier, the occupational employment shift index is the difference of the occupational employment ratios (employment of one substitutable occupation in one manufacturing industry to the employment of that occupation in the entire economy) between two time points. The change rate of the occupational employment shift index is the proportion of this difference to the base year ratio. The formula for calculating the average annual change rate of the occupational employment shift index related to the substitutable occupation was the same as (3.6) except substituting occupational employment ratios for occupational matrix ratios in the formula.
The data in the form of occupational employment available from the BLS (1978) do not cover exactly the same time period as the one in the study. Only 1970 and 1976 data were published. Hence, this average annual change rate was only an approximate estimate of the true value. If the assumption of constant change rate holds, this estimation should be reasonably close to the true value.

3.3.5 Data for Sectoral Growth

For the variable of sectoral growth, the average annual change rates of industry output and the average annual change rates of capital/labor ratio in industries were used as the indicators. These change rates were distributed to substitutable occupations in proportion to their employment to the total employment of the industry to which the substitutable occupation belongs. Then the weighted average of the change rates according to the employment proportions was used to represent the change rates of substitutable occupations in relation to a new and emerging occupation.

The formula for calculating their respective change rates is as follows:

\[
U_j = \frac{\sum_{k} E_k \sum_{k} E_k}{\sum_{k} E_k} \left( \sum_{i} U_i E_{ki}/E_i \right) \tag{3.7}
\]

\[
C_j = \frac{\sum_{k} E_k \sum_{i} C_{ki}}{\sum_{k} E_k} \left( \sum_{i} C_{ki}/E_i \right) \tag{3.8}
\]

where \( U_j \): average annual change rate of industry output in connection to new and emerging occupation \( j \)

\( U_i \): average annual change rate of output of industry \( i \) to which the occupation substitutable for new and emerging occupation \( j \) belongs;
C_j : average annual change rate of capital/labor ratio in connection to new and emerging occupation j

C_i : average annual change rate of capital/labor ratio of industry i to which the occupation substitutable for new and emerging occupation j belongs

E_k : employment of substitutable occupation k

E_{ki} : employment of substitutable occupation k in industry i

E_i : employment of industry i

m : number of substitutable occupations for new and emerging occupation j

n : number of industries within which the substitutable occupations are distributed

The data set used for calculating U_i was from the Time Series Data for Input-Output Industries (1958-1976), published by the Bureau of Labor Statistics (1979). These data were price-deflated and the change rates were obtained by using least squares method. Thus, only minor transformation was needed to obtain the data from U_i.

The data set used for calculating C_i was from the Total Factor Productivity data series published by Kendrick and Grossman (1980) from 1948 to 1976. The least squares ratio or heuristic ratio, whichever appeared better, was computed and adopted.

3.3.6 Data for Technological Change

For the variable of technological change, two indicators were suggested. One was the average annual change rate of total factor productivity and the other was the average annual change rate of patent intensity ratio. Since the Kendrick's TFP index was preferred, the average annual change rate of Kendrick's TFP was calculated.
The time series data prepared by Kendrick and Grossman (1980) on Kendrick's TFP from 1948 to 1976 with respect to two-digit SIC codes were used. However, the TFPs were aggregated at two-digit SIC level. Their distributions to the substitutable occupations in proportion to the employment of the substitutable occupations to their respective industry employment could be calculated using the following formula:

\[ K_j = \frac{\sum_k (E_k/E_k)}{\sum_i (E_i/k_i/E_i)} \]

(3.9)

where 

- \( K_j \): average annual change rate of Kendrick's TFP in connection to new and emerging occupation \( j \)
- \( E_k \): employment of substitutable occupation \( k \)
- \( E_{ki} \): employment of substitutable occupation \( k \) being distributed in industry \( i \)
- \( E_i \): employment of industry \( i \)
- \( k_i \): average annual change rate of Kendrick's TFP of industry \( i \)
- \( m \): number of substitutable occupations for new and emerging occupation \( j \)
- \( n \): number of industries within which the substitutable occupations are distributed

For the average annual change rate of patent intensity ratio, similar procedure was taken. The patent statistics from the Office of Technology Assessment and Forecast, United States Patent and Trademark Office, were used to calculate the patent intensity ratios of respective industries (number of patents granted in a manufacturing industry to the total number of patents granted in the entire manufacturing industry in one year) during the time period of concern. The average annual change rate could be determined using
either least squares or heuristic approach, as discussed earlier. Since the change rates were aggregated at two-digit SIC level, their distributions to the substitutable occupations could be calculated using the same formula as (3.8). Only \( K_j \) and \( K_i \) were substituted by \( P_j \) and \( P_i \) respectively. Here \( P_j \) represents the average annual change rate of patent intensity ratio in relation to new and emerging occupation \( j \). \( P_i \) represents average annual change rate of patent intensity ratio of industry \( i \). Missing data were assigned to the substitutable occupations with no patent statistics available to their corresponding industries.

3.4 DATA FOR EXOGENOUS VARIABLES

In this section, two exogenous variables are introduced. The sources for the variables of research and development investment in government and private companies are treated.

3.4.1 Data for Research and Development

For research and development which contributes directly to technological change, two variables were suggested. One was the average annual change rate of government-financed R & D intensity ratio in industries, while the other was the average annual change rate of privately financed R & D intensity ratio in industries. These intensity ratios are the proportion of annual R & D expenditure in industries to their respective annual sales. The data for annual R & D expenditures in two-digit SIC manufacturing industries were obtained from the biannual report of Science Indicator (National
Science Board, 1977). The percentages of these annual R & D expenditures distributed by either federal government or private companies were given in the annual report of Research and Development in Industry by the National Science Foundation. The multiplication of these two could produce the annual R & D expenditures in manufacturing being distributed by government and private companies respectively. The data for annual sales in manufacturing were obtained from the 1977 Statistical Supplemental to the Survey of Current Business (Department of Commerce, 1977). Then, the average annual change rates of R & D intensity ratio in industries could be calculated either by least squares or heuristic approach.

These average annual change rates of R & D intensity ratio in industries were redistributed to the substitutable occupations in relation to the new and emerging occupation using formula (3.9) only substituting $R_{ijg}'$, $R_{ijp}'$, and $R_{ijg}$, $R_{ijp}$ for $K_j$ and $K_i$, respectively. Here $R_{ijg}$ represents the average annual change rate of R & D intensity ratio in government in relation to the new and emerging occupation $j$; $R_{ijp}$ represents the average annual change rate of R & D intensity ratio in private companies in connection to new and emerging occupation $j$; $R_{ig}$ represents the average annual change rate of R & D intensity ratio in government in industry $i$; and $R_{ip}$ represents the average annual change rate of R & D intensity ratio in private companies in industry $i$. Similarly, for industries with which no R & D expenditure data were available, missing data were assigned to their corresponding substitutable occupations.
3.5 **SUMMARY**

In summary, this chapter has filled partially the empirical part of the model. The population, sample, selection of substitutable occupations, sources of data for each indicator, and data transformations were described in this chapter.
Chapter IV

METHODOLOGICAL CONSIDERATIONS

The purpose of this study was to explore causal relationships among variables, which were selected according to existing theories and evidences, in an attempt to explain how the changes in distributions of new and emerging occupations appropriate for T & I training have been developed in manufacturing. It appeared that several words concerning the logic of causal inference needed to be asserted before displaying the methodology.

It is alerted that causality in social science is more difficult to establish than it is in natural and physical science. This difficulty is due to the fact that researchers in social science are in general unable to control adequately for intervening variables. In economics, the meanings of causation (explicit causal chains vs. interdependent system) and the ways to establish it (simple vs. simultaneous system) are still controversial (L'esperance, 1972). In sociology, major efforts have been launched toward social causal orderings and further theory construction about social process under non-experimental conditions (Blalock, 1964). In psychology, researchers are more fortunate than their colleagues in other disciplines in establishing causal orderings because experimental controls are sometimes feasible.
It was presented by Kerlinger and Pedhazur (1974) that prediction need not have a sound understanding of the causal relationships between dependent and independent variables, especially when using multiple regression to fit the data. It is true when the purpose of the study is to obtain a pattern of empirical association among variables, a prior theory governing the behaviors is not absolutely essential. However, when explanation of a phenomenon or behavior is desired, knowledge of causal orderings of the variables involved becomes crucial to the power of explanation. Explanation here means finding the determining factors and conditions for the known event. It was argued here that even for the purpose of prediction (here means determining the effects given initial conditions), a better understanding of the interactions of the variables would enhance the predictive validity of the model. Therefore, explanation should take precedent step in any attempt to predict.

In order to explain the factors contributing to the phenomenon, related theories concerning the phenomenon should be organized and causal orderings of the chosen variables should be established. In the meantime, hypotheses relating to causal inference of the variables have to be formulated and tested.

Based on the diagram in Figure 2, a non-recursive causal model with the changes in distributions of T & I new and emerging occupations as the endogenous variable was to be explained by the structural relationships of the intervening endogenous and predetermined
exogenous variables. Hence, the method for analysis lay within the general framework of causal analysis.

Causal analysis or modeling has its origin in the fields of genetics and economics, and recently has been explored extensively in sociological research. It is a method to link theory and empirical data together in order to determine the causal orderings of a phenomenon. Theories are used to establish possible cause-and-effect relationships among variables of concern. Empirical data are subsequently used to verify and then modify the causal specification. The model involves a set of equations which hypothesizes causal relations among variables that comprise the phenomenon or behavior under investigation. In sociology, it is called path analysis. In econometrics, it is called simultaneous equations system. However, a more generic term, structural equation model, is preferred since the causal mechanisms are embodied in the structure of an equation system which generates the observable variables.

In a structural equation model, four developmental steps—specification, identification, estimation, and hypothesis testing—deserve individual attention. They are treated in the following sections.

4.1 SPECIFICATION

Heise (1969) argued strongly that theory is absolutely essential to formulating any causal model. He regarded it as an assumption that the causal laws governing the system are established
sufficiently to specify the causal priorities among variables in a way that is "undebatable." In other words, the specification supported by existing theories takes the primary consideration in model formulation. It was not argued here the importance of theory in connection to the validity of the model. But what was argued was that in an exploratory study, like the one under pursuit, there might not be enough convincingly established theories or empirical evidence to specify all the causal priorities. Thus, an exploratory attempt to specification along with more emphasis on empirical verification was to be taken here.

The specification of a structural equation model is dependent upon the proposed ordering of the behaviors. It may be a recursive or a non-recursive one, or it may be associated with observable or unobservable variables. For various model specifications in sociological examples, readers may wish to review research done by Wiley and Hauser (1977).

For the present study, the model proposed was apparently a non-recursive one. Both the path analytic method and simultaneous equation system are appropriate for serving as the methodological framework in model building. However, path analytic was preferred here for two reasons. First, under path analytic method, the causal orderings are specified explicitly while in simultaneous equation system, interdependence rather than causation is emphasized. Because causality was desired in this study, path model was more appropriate. Second, and perhaps most important, was that most of the variables (i.e., technological change, occupational structural change,
occupational demand and supply effects) in the model were abstract concepts or theoretical constructs (here abstract concept, theoretical construct, unobservable variable, unmeasured variable, and latent variable are used interchangeably throughout the text) which were not directly measurable, or even if measured, substantial errors were anticipated. It is perfectly all right that abstract concepts or theoretical constructs which represent theories or hypotheses formulated are not directly measurable because it may indicate poorly formulated hypotheses or theories in terms of its scope and generality if otherwise measured (Sullivan and Feldman, 1979). However, the crucial thing was how to test hypotheses based on abstract concepts if they were not directly measured. Therefore, ways of measuring these concepts or constructs and of dealing with measurement errors were essential to the problem of concern. Simultaneous equation system did not provide much help, especially for measurement error, although the problem has been given attention recently (Geraci, 1976). However, dealing with measurement error in causal model has been a major effort in the development of sociological methodology (Blalock, 1968, 1969a, 1969b, 1971; Costner, 1969; Duncan, 1975; Mayer and Younger, 1975; Siegel and Hodge, 1968; Sullivan, 1974; Wert, Joreskog, and Linn, 1973; Wiley, 1973; Zeller and Carmines, 1980) in the last decade or longer. Thus, recent achievements in these developments can definitely shed some light on this problem of measurement error.

Jacobson and Lalu (1974) suggested three alternative techniques: single indicator, composite index, and multiple indicators, which can be employed in measuring theoretical constructs. There are certain
problems associated with using single indicator or composite index. Among the most serious problems is the specification error (non-random measurement error). In single indicator situations, the assumption of no specification error has to be maintained in order to obtain unbiased parameter estimation because it is impossible to detect any specification error using single indicator. The same problem applies to composite index, also. In addition, single indicator still suffers from the assumption of accounting for 100 percent of variation in the variable by the single indicator. Similarly, composite index suffers from: (1) no unique measure since the measure depends upon the number of items included in the index, the weights assigned to particular items, and the manner in which the items are combined, and (2) lack of theory in the interpretation of the combined variable once combined.

Multiple indicators, on the other hand, have certain advantages over the other two. There are conceptual and statistical advantages associated with using multiple indicators to represent theoretical construct. First, from the conceptual standpoint, more indicators can tap more dimensions of the variables and thus be more representative. Second, from the statistical standpoint, the more indicators the unmeasured variable has, the more variance of the unmeasured variable can be explained. Costner (1969) suggested at least three gains from using multiple indicator approach: (1) it is possible to obtain an estimate for each of the specified unknown parameters (coefficient) even if the system of equation is over-identified; (2) it is possible to detect specification error, and (3) it is possible to test the
implications of the causal model outlined in the "main" theory (as contrasted to the "auxiliary theory" discussed later). However, multiple indicators approach is not free from difficulties and problems. These problems were treated a little later in the study.

Granted that multiple indicators approach is better than others in terms of handling measurement of theoretical constructs and problem of measurement error then the questions become what indicators should be included and how many? Blalock (1968) called for "auxiliary theory," which is to specify the relationships between theoretical and empirical worlds, or more precisely, between theoretical constructs and empirical indicators to handle measurement problems. These relationships have been called in various ways; for instance, epistemic correlations, rules of correspondence, and operational definition. Statistically speaking, the square of the epistemic correlation coefficient is equal to the proportion of variance in the indicator accounted for by the concept. The standardized correlation coefficient corresponds to the reliability coefficient which indicates the extent to which random measurement error occurs. However, it is much more difficult to obtain a quantity indicating the degree of validity (caused by non-random measurement error) of the indicator to the theoretical construct. In other words, determining non-random measurement error (means to determine the extent to which systematic errors other than theoretical construct affecting the indicators) is much more difficult. Additional information is needed for this attempt.
The number of indicators selected required some conceptual and statistical considerations. It was mentioned earlier that more indicators have better representation and variance explanation. However, this can not be generalized without further qualification. Sullivan (1974) indicated firstly that in a recursive model, the intraset correlations among indicators of the latent variables must be as similar as possible and highly correlated in order to avoid what Gordon (1968) called unequal redundancy. Here redundancy means the correlations among indicators of a variable. Unequal redundancy means the correlations among indicators on one variable are not equal to those of the others. Consequently, the regression coefficients for indicators with low correlations are more likely to achieve significance. Secondly, the number of indicators in each latent variable should be as equal as possible in order to avoid the "differential repetitiveness." In regression, the total effect of an independent variable on the dependent variable is to be divided among its indicators. The larger the number of indicators, the smaller effect each indicator has on the dependent variable. Then it is more likely that none of the indicators will achieve significance. But for a situation of weaker independent variable with fewer indicators, each indicator may achieve significance since the effects are "concentrating" rather than "spreading out" by a number of indicators. In other words, equal number of indicators in each variable would at least assure their equal chance of contribution to variance explanation. Furthermore, the statistical problem of under-identification may surface if
the number of indicators are unequal and single indicator is used as the measure of a variable (Blalock, 1969).

Summarizing the discussion, one can now present the criteria for indicators selection in the following order of importance: (1) there must exist a theoretical relationship between latent variables and their indicators; (2) intraset correlations among indicators of a latent variable must be high and as similar as possible to those of other variables; and (3) there must be an equal number of indicators per variable. Nevertheless, these are guidelines for selection and it may be unrealistic to expect that each indicator exactly follows the rules. As a matter of fact, there are few studies reported following all guidelines described above (Sullivan, 1974).

Very often, the indicator will be rejected based not on the above guidelines, but on its detection of non-random measurement error. Costner (1969) presented four possible sources of non-random measurement error due to various spurious correlations among indicators. Some of them are detectable and some are not depending upon how many indicators are included in each construct. Zeller and Carmines (1980) demonstrated that an indicator can be rejected if it is the source of non-random measurement error.

Following discussion of the characteristics of multiple indicators and their selection in a causal model, it appeared that specifying the model in structural equations would be helpful for later consideration. Thus, the model in Figure 3 is elaborated in Figure 4, in which the symbols used to represent the variables, the indicators, and their residuals of error terms are included.
Figure 4: Multiple Indicators Causal Model with Symbols and Error Terms

The symbols in Figure 4 mean:

$X_i$: exogenous measured variables

$Z_i$: endogenous latent variables

$Y_i$: indicator for endogenous variables

$a_i$: path coefficients of the latent structure

$b_i$: epistemic correlation coefficients

$e_i$: error terms for indicators of endogenous variables

$r_i$: residuals of endogenous variables

$s_i$: residuals of exogenous variables
The indicators in Figure 4 were selected based mainly on auxiliary theories in the section of substantive considerations. It could also be noted from the structural equations that two types of structural equations, the structure of latent variables and the structure of measurement variables, were specified.

The structural equations describing the latent structure given in Figure 4 are as follows:

\[ Z_1 = a_1 X_1 + a_2 X_2 + a_5 Z_2 + r_1 \] (4.1a)
\[ Z_2 = a_4 Z_1 + r_2 \] (4.1b)
\[ Z_3 = a_3 Z_1 + a_6 Z_2 + r_3 \] (4.1c)
\[ Z_4 = a_7 Z_3 + a_{10} Z_5 + r_4 \] (4.1d)
\[ Z_5 = a_9 Z_4 + r_5 \] (4.1e)
\[ Z_6 = a_8 Z_3 + a_{11} Z_4 + a_{12} Z_5 + r_6 \] (4.1f)

The structural equations describing the measurement model in the Figure are as follows:

\[ Y_1 = b_1 Z_1 + e_1 \] (4.2a)
\[ Y_2 = b_2 Z_1 + e_2 \] (4.2b)
\[ Y_3 = b_3 Z_2 + e_3 \] (4.2c)
\[ Y_4 = b_4 Z_2 + e_4 \] (4.2d)
\[ Y_5 = b_5 Z_3 + e_5 \] (4.2e)
\[ Y_6 = b_6 Z_3 + e_6 \] (4.2f)
\[ Y_7 = b_7 Z_4 + e_7 \] (4.2g)
\[ Y_8 = b_8 Z_4 + e_8 \] (4.2h)
\[ Y_9 = b_9 Z_5 + e_9 \] (4.2i)
\[ y_{10} = b_{10}z_5 + e_{10} \]  \hspace{1cm} (4.2j)

\[ y_{11} = b_{11}z_6 + e_{11} \]  \hspace{1cm} (4.2k)

4.2 IDENTIFICATION

Following the discussion of the multiple indicators approach to model specification with abstract concepts and the problems associated with indicators selection, the problem of identification needed immediate examination. Heise (1969) stated that identification problem is to estimate unknown parameters in a model from available empirical data. These unknown parameters, so-called path coefficients, are estimated through correlations or covariances between pairs of variables. These estimations are done by setting up systems of equations expressing each empirical correlation as a function of path coefficients. Algebraically speaking, identification is whether there is sufficient information included in the system of equations to solve for unknown parameters.

Asher (1976) indicated that in a recursive system, two restrictions called coefficient and covariance restrictions are generally imposed in order to guarantee that the equations will be identified. Coefficient restriction is that the causal orders between variables have been assumed unidirectional, which eliminates almost half of the parameters to be estimated if the causal orders are assumed otherwise. Covariance restriction is zero correlations of pairs of disturbance terms.

In a non-recursive system, two coefficient restrictions frequently discussed in econometrics texts need to be considered.
The first one is the order condition. It means that in a structural equation model consisting of \( k \) linear equations, it must exclude at least \( k-1 \) variables in the equation in order for any equation to be identified. It can be explained more clearly by using the model in Figure 4 as an example. First, the coefficients of each latent variable in the structural equation system (excluding the measurement structure for the time being) can be expressed in matrix form as follows:

\[
\begin{bmatrix}
 a_1 & a_2 & 0 & a_5 & 0 & 0 & 0 & 0 \\
 0 & 0 & a_4 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & a_3 & a_6 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & a_7 & 0 & a_{10} & 0 \\
 0 & 0 & 0 & 0 & 0 & a_9 & 0 & 0 \\
 0 & 0 & 0 & 0 & a_8 & a_{11} & a_{12} & 0
\end{bmatrix}
\]

Since there were six (\( k \)) linear equations, five (\( k-1 \)) coefficients in each equation should be zeros for that equation to be identified. Examining the coefficients matrix, it was found that all of them satisfied this condition. In addition, some of them had more than five zero coefficients, an indication of over-identification.

However, the order condition is a necessary but not a sufficient condition for identification. This means that if the order condition is not met, the equation will not be identified. But if the condition is met, it is still possible to have a model in which some equations
are under-identified, some exactly identified, or some over-identified (Christ, 1966). Therefore, a second condition which is the rank condition is needed.

The rank condition means that an equation in a model of k linear equations is identified if and only if at least one non-zero determinant of k-1 rows and columns is contained in the coefficient matrix after all columns of coefficients not having a zero entry and the rows of coefficients of that equation are omitted. If more than one non-zero determinant is found, the equation is over-identified. The rank conditions can be demonstrated by using the model at hand. It should, in each equation, have at least one 5 x 5 non-zero matrix after taking the non-zero columns and rows out of the matrix. Since more than one 5 x 5 non-zero matrices could be obtained in the equations, it indicated clearly that over-identification had occurred, except for the first and last ones.

The identification problem discussed so far focused only on the structural model of latent variables. The identification of the measurement model is not included. In order for the parameters in the measurement model to be identified, the order and rank conditions must hold for the latent structure in the first place. The identification for parameters of the measurement model can be pursued under a general model which is the combination of latent and measurement structures. It was suggested that if the matrix of partial derivatives of the latent parameters with respect to measurement parameters has rank equal to the number of parameters in the measurement structure to be estimated, then the parameters of the measurement structures
are identified (Wiley, 1973). The key to the problem is the specifica-
tion of the general model which involves the analysis of covariance
structure, a topic most appropriately treated in the section of
parameters estimation. Hence, the problem of identification for
measurement parameters was treated later.

Regardless of the identifiability of measurement parameters,
the identification of latent parameters should take the higher
consideration. It was indicated earlier that the model of latent
structure in the study was over-identified. In other words, there
would not be a unique solution for each parameter estimated.
Furthermore, if the measurement parameters were also just-identified
or over-identified (the chances are that they would be), excessive
information for parameters became a problem. Then, how to use this
additional information effectively became a crucial point to the
validity of the parameters to be estimated. Costner (1969)
demonstrated that this additional information can be used to test the
consistency of the estimated parameters and to detect non-random
measurement error. Duncan (1975) also indicated similar gains from
over-identification. Nevertheless, these gains from additional
information are embedded in the method of estimation. Hence, it led
to the discussion of parameter estimation, the topic treated in the
next section.

4.3 ESTIMATION

In a structural equation system with directly observable
variables, parameters estimation for an over-identified system of
equations can be done by two-stage least square (2SLS), a method found frequently in econometrics. However, in a multiple indicators model, when over-identification occurs, there will be multiple estimates of the same parameter. If the model is correctly specified, the difference among multiple estimates of the same parameter will be due only to sampling errors. Consistency criteria indicated earlier can be tested for significant difference. Sullivan and Feldman (1979) proposed a formula to compute the t statistics between pairs of consistency criteria. If significant differences occur, model revision is warranted.

Suppose that all tests fail to achieve significance and there is no substantive problem associated with the model. The problem of multiple estimates of each parameter still remains despite the fact that there may indeed be little difference. Blalock (1969) suggested taking the average of each estimate. Duncan (1975) proposed averaging the estimation equations to produce a single value. Regardless of the method chosen, there will be unbiased estimates of the population parameters.

Unbiasedness is certainly a desirable property of an estimator. However, it may not be the most crucial one. In general, an estimator should at least have the following three properties. First, an estimator is unbiased in the sense that if \( X \) is an unknown population parameter and \( x \) is an estimate of \( X \), the expected value of \( x \) should be equal to \( X \). Second, an estimator is efficient in the sense that when sample size approaches infinity, \( x \) should be equal to \( X \) (Chu, 1972; Wonnacott and Wonnacott, 1970). Consistency is a rather
weak property since it is common sense in statistics that the larger the sample becomes, the closer is the sample estimate to population parameter. However, unbiasedness and efficiency are rather important properties for any estimator to be valid and reliable.

Costner's path analytic approach to parameter estimation in multiple indicators model was discussed earlier. And it is recognized that under the condition of over-identification, unbiased multiple estimates are obtained. Averaging these multiple estimates will lead to a single unbiased estimate but not an efficient one. The reason for being inefficient is because the average estimate does not reflect variations in individual estimates and consequently, results in greater sampling variability. Hauser and Goldberger (1971) demonstrated a more efficient estimating procedure using maximum likelihood approach.

Maximum likelihood estimation grew out of the work of Lawley (1943) in factor analysis. Factor analysis is to study the variation and covariation (or correlation) in a set of observed variables being explained by a smaller number of unobserved factors (common and unique factors). Each variable is conceptually a linear function of one or more factors. The analysis of causal model with unobservable variables represented by multiple indicators can be put under the framework of confirmatory factor analysis. The difference between confirmatory and exploratory analysis lies in the fact that indeterminacy in the confirmatory factor model is eliminated in a more natural way by constraining the values of a number of elements in the
factor loading matrix, the factor covariance matrix, and the error variance matrix. These constraints stem from prior knowledge of the phenomenon being studied and this prior knowledge is incorporated into model specification. In general, the particular element of factor loading and covariance matrices are specified zeros. However, in exploratory factor analysis, the elements in all the matrices are free to estimate. Putting the causal model and factor model together, the latent variables, the indicators, and their associated errors in the causal model can be viewed as unobserved common factors, the observed variables, and the unique factor respectively in the factor model. Then, the maximum likelihood estimation procedure appropriate for factor analysis is also appropriate for causal analysis although their interpretations should be different (Goldberger, 1973).

Conceptually speaking, the maximum likelihood estimator (\( \hat{\Theta} \)) of the population parameter (\( \Theta \)) is that the value \( \hat{\Theta} \) for which the associated likelihood function is maximized. A likelihood function is the density function generating the observed values of the variables of some known form.

Long (1976) suggested four advantages of maximum likelihood estimation (MIKE) procedure over the heuristic one of path analysis in over-identification situation. First, unlike path analytic procedure where only the information sufficient to obtain the estimates are used, all the information contained in the model are used simultaneously in order to generate the best estimate. Second, the estimates obtained are scale-invariant. This can facilitate the use of standardized path coefficient. In addition, when the units
of measurement have no intrinsic meaning, correlation matrix may be used in analysis instead of covariance matrix. Third, these estimates are consistent, efficient, and asymptotically normal and unbiased. That means when the sample size approaches infinity, the estimates are normally distributed and unbiased. Fourth, a likelihood ratio test is provided to test the goodness of fit of the model under the null hypothesis. If poor fit of parameters to the data occurs, this likelihood ratio test will provide a clue to any misspecification of the model. However, this test can only tell whether the model fits but not where the model does not fit.

As a matter of fact, confirmatory factor analysis model is one of the special cases of the general model of analysis of covariance structures (Joreskog, 1970; Joreskog and Sorbom, 1978; Long, 1976; Werts, Joreskog, and Linn, 1973). The analysis of covariance structures model is to handle a set of linear structural equations in which either latent or observed variables are included. It allows the model to contain errors in both the equations (residuals, disturbance) and the observed variables (measurement errors). The basic equations determining two models, the structural model and the measurement model, are expressed in the following:

\[
\text{Structural model} \quad B\xi = \Gamma\zeta + \xi \quad (4.3)
\]

\[
\text{Measurement model} \quad x = \Lambda_x \eta + \epsilon \quad (4.4)
\]
\[
\quad y = \Lambda_y \xi + \delta \quad (4.5)
\]
where \( \eta : mxl \) vector of true latent dependent variables
\[ \xi : nxl \] vector of true latent independent variables
\[ \zeta : mxl \] vector of errors in equations
\[ B : mxm \] coefficient matrix for true dependent variables
\[ \Gamma : mxn \] coefficient matrix for true independent variables
\[ y : pxl \] vector of observed dependent
\[ x : qxI \] vector of observed independent variables
\[ \Lambda_y : pxm \] regression matrix of \( y \) on \( \eta \)
\[ \Lambda_x : qxn \] regression matrix of \( x \) on \( \xi \)
\[ \epsilon : pxl \] vector of errors of measurement in dependent variables
\[ \delta : qxI \] vector of errors of measurement in independent variables

The structural model specifies the causal relationships, either recursive or non-recursive, among latent variables. The measurement model specifies how the latent variables are measured by their observed indicators.

In the structural model, it is assumed that:

\[ E(\eta) = E(\xi) = 0 \] \hspace{1cm} (4.6a)
\[ E(\xi) = 0 \] \hspace{1cm} (4.6b)
\[ E(\xi \xi') = 0 \] \hspace{1cm} (4.6c)

and \( B^{-1} \) exists. It is also noted that in equation (4.3) if \( B \) is an identity matrix, it can be thought of as a vector of observed variables and \( \xi \) and \( \xi \) are vectors of common and unique factors, respectively, the same as the factor model described earlier.
In the measurement model, it is assumed that:

\[ E(\varepsilon \eta') = 0 \]  \hspace{1cm} (4.7a)
\[ E(\varepsilon \xi') = 0 \]  \hspace{1cm} (4.7b)
\[ E(\eta \xi') = 0 \]  \hspace{1cm} (4.7c)
\[ E(\delta \varepsilon') = 0 \]  \hspace{1cm} (4.7d)
\[ E(\delta \xi') = 0 \]  \hspace{1cm} (4.7e)
\[ E(\theta \zeta') = 0 \]  \hspace{1cm} (4.7f)

In other words, the residuals in the measurement model are not correlated with the true latent variable and the residuals in the structural model. However, it is noted that \( \varepsilon \) and \( \delta \) may be correlated.

Let \( \Phi(n \times n) \), \( \Psi(m \times m) \), \( \Theta_{\varepsilon}(p \times p) \), \( \Theta_{\delta}(q \times q) \) be the covariance matrices of \( \xi, \xi', \varepsilon, \) and \( \delta \) respectively. The covariance matrix of \( \eta \) can be expressed as follows:

\[
\text{Cov}(\eta) = E(\eta \eta') = E[(B^{-1} \Gamma \varepsilon + B^{-1})(B^{-1} \Gamma \varepsilon + B^{-1} \xi')]
\]

\[
= B^{-1}(\Gamma \Phi \Gamma^\prime + \Psi)B^{-1}
\]  \hspace{1cm} (4.8)

Then, the covariance matrix \( (\Sigma) \) of the observed data \( Z' = (y', x') \) can be determined as follows:

\[
\Sigma = E\begin{bmatrix}
\{Y' \quad Y'^2 \}
\{X \quad X^2 \}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\Lambda_y B^{-1}(\Gamma \Phi \Gamma^\prime + \Psi)B^{-1} \Lambda_y^\prime + \Theta_{\varepsilon}^2 & \Lambda_y B^{-1} \Gamma \Phi \Lambda_x^\prime \\
\Lambda_x \Phi \Gamma^\prime B^{-1} \Lambda_y^\prime & \Lambda_x \Phi \Lambda_x^\prime + \Theta_{\delta}^2
\end{bmatrix}
\]  \hspace{1cm} (4.9)

It can be seen from equation (4.9) that the elements in \( \Sigma \) are functions of elements in eight matrices, namely \( \Lambda_y \), \( \Lambda_x \), \( B \), \( \Gamma \), \( \Phi \), \( \Psi \), \( \Theta_{\varepsilon} \), and \( \Theta_{\delta} \). Since parameters in equations (4.3), (4.4), and (4.5)
can not be estimated directly, equation (4.9) is employed in practice for parameters estimation.

The estimation procedures, the same as those in the confirmatory factor analytic model described above, are to fit the $\Sigma$ imposed by the model to the sample matrix $S$. Here, the elements in $\Sigma$ are imposed by specifying whether the elements in those eight matrices are fixed, constrained, or free. Since the elements in the matrices are the parameters to be estimated, fixed parameters mean that the parameters have been assigned values; constrained parameters mean that equal value of one or more other parameters has been given to a parameter; and free parameters mean that the parameters are constrained to any value and are free to be estimated. These specifications are done according to model specifications indicated earlier in this section. Equations (4.1) and (4.2) are equivalent to equations (4.3) and (4.5). They are shown in matrix form in equations (4.10) and (4.11), respectively, on the next page. It should be noted that equation (4.4) was not used here since all the exogenous variables are measured variables.

Before getting into the details of estimation, attention is called to the problem of identification of the general analysis of covariance structures. It was discussed earlier that the structural model was identified according to the rank and order conditions suggested in econometrics texts where no measurement error is assumed. When the measurement errors are explicitly incorporated into the model, the identifiability of the model means whether all parameters of the model are identified, which in turn depends on the specification
The structural equations are:

\[
\begin{bmatrix}
1 & -a_5 & 0 & 0 & 0 & 0 \\
-a_4 & 1 & 0 & 0 & 0 & 0 \\
-a_3 & -a_6 & 1 & 0 & 0 & 0 \\
0 & 0 & -a_7 & 1 & -a_{10} & 0 \\
0 & 0 & 0 & -a_9 & 1 & 0 \\
0 & 0 & -a_8 & -a_{11} & -a_{12} & 1
\end{bmatrix}
\begin{bmatrix}
Z_1 \\
Z_2 \\
Z_3 \\
Z_4 \\
Z_5 \\
Z_6
\end{bmatrix}
= \begin{bmatrix}
a_1 \\
a_2 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
r_1 \\
r_2 \\
r_3 \\
r_4
\end{bmatrix}
\]

The measurement equations are:

\[
\begin{bmatrix}
Y_1 \\
Y_2 \\
Y_3 \\
Y_4 \\
Y_5 \\
Y_6 \\
Y_7 \\
Y_8 \\
Y_9 \\
Y_{10} \\
Y_{11}
\end{bmatrix}
= \begin{bmatrix}
b_1 \\
b_1 \\
b_3 \\
b_4 \\
b_5 \\
b_6 \\
b_7 \\
b_8 \\
b_9 \\
b_{10} \\
b_{11}
\end{bmatrix}
\begin{bmatrix}
Z_1 \\
Z_2 \\
Z_3 \\
Z_4 \\
Z_5 \\
Z_6 \\
Z_7 \\
Z_8 \\
Z_9 \\
Z_{10} \\
Z_{11}
\end{bmatrix}
+ \begin{bmatrix}
e_1 \\
e_2 \\
e_3 \\
e_4 \\
e_5 \\
e_6 \\
e_7 \\
e_8 \\
e_9 \\
e_{10} \\
e_{11}
\end{bmatrix}
\]

of fixed, constrained, and free parameters in the model. If each parameter of the model can be expressed in terms of the covariance or correlation (in the case of standardized variables) of the observed variables, the model is identified. Although some sufficient
conditions for identifiability were given by Wiley (1973), it is quite difficult to do the computation by hand. Fortunately, there is a built-in step in the estimation procedure in a canned program called LISREL to check the identification status of the model. This is done by computing the information matrix for all independent unknown parameters at the beginning of the iterative estimation. If the information matrix is non-singular, the standard errors of all estimated parameters can be computed and the model is identified (Joreskog and Sorbom, 1978). If the parameters are not identified, printed message will be sent out by the program. Despite the capability of the computer program in detecting the status of identification, it seems that not until some sort of sufficient and necessary conditions (analogous to the rank and order conditions in econometrics) for identification can be established, the problem of identification still remains troublesome in the use of analysis of covariance structures in causal modeling (Long, 1976).

Having discussed the problem of identification, we return to the root of the matter, which is the estimation. The estimation problem, as indicated earlier, is essentially to fit $\Sigma$ to the sample matrix $S$ with the fitting function being minimized with respect to independent parameters. The fitting function is:

$$F = \log|\Sigma| + \text{tr}(\Sigma^{-1}) - \log|S| - (p+q)$$

(4.12)

where $|\Sigma| :$ determinant of $\Sigma$

$|S| :$ determinant of $S$

$\text{tr} :$ trace of the matrix
q : number of observed exogenous variables
p : number of observed endogenous variables

Under the assumption of multinormality of those observed variables, minimization of the function employs iterative procedure which converges from an arbitrary starting point to a minimum of F. And the maximum likelihood estimates of the parameters are obtained.

When numerical analysis method is involved in the data analysis, computer program becomes a must. The most flexible canned program dealing with general analysis of covariance structures and using maximum likelihood esitmate is LISREL. A new version of this program, LISREL IV, released in 1978, should suit the estimation at hand. The details of this program were treated in a later chapter of data analysis.

4.4 MODEL TESTING

After the parameters in the model are estimated, the next step is to examine the adequacy of the structure of the model. In other words, the goodness of fit of the model to the data needs to be tested. Burt (1973) suggested that internal and external validity be used as the criteria for model adequacy. Internal validity is concerned with parsimonious representation of the complete set of observed variables with a small number of latent variables. External validity is concerned with the ability of the structure to make prediction in comparison to other structures specifying different
causal relationships. For internal validity, to gain the advantages of using maximum likelihood estimate, a likelihood ratio test is available.

Let $H_0$ be the null hypothesis of the model under given specifications and $H_1$ be the alternative hypothesis for the least restrictive model specification (more free parameters to be estimated and, hence, fewer assumptions about the model). Let $F_0$ be the minimum of $F$ (the fitting function) under $H_0$ and $F_1$ the minimum of $F$ under $H_1$. Then minus twice the logarithm of the likelihood ratio becomes $(N/2)(F_0-F_1)$, which is distributed approximately as $x^2$ with degrees of freedom equal to the difference in the number of independent parameters estimated under $H_0$ and $H_1$. The degrees of freedom under any hypothesis is equal to $\frac{1}{2}(p+q)(p+q+1)-t$, where $t$ is the total number of independent parameters estimated under the hypothesis (Joreskog and Sorbom, 1978).

A major advantage of this likelihood ratio test is to allow successive model testing or fitting while relaxing model assumptions and adding independent parameters successively. It is particularly appropriate for studies of a more exploratory nature, like the one undertaken. However, these successive tests or fitting can only be valid if the successive models are nested. Here nested model means that a model can be obtained from the other by constraining one or more of the free parameters in the later model to be fixed. Then the next question is when to stop this sequential fitting. Up to the present time, there is no specific objective criterion that can guide the decision. A general rule is that when the $x^2$ value drops to a considerable extent compared to the difference in degrees of freedom,
it indicates a real improvement. If the $x^2$ value starts remaining stable, it indicates that added parameters may not have significant meaning. Another way to help check the fitting is to inspect the discrepancies between the observed matrix $S$ and the estimated matrix $\Sigma$. If there is significant discrepancy, the associated observed variable should be examined. However, a word of caution in this comparison is that it may be misleading in certain cases as demonstrated by Costner and Schoenberg (1973).

While sequential fitting of different models is somewhat subjective, one should bear in mind that statistical testing can only be meaningful if it is guided by established theories. Therefore, both the theoretical and statistical significance should be considered when revising the model. In addition, the practical consideration in model revision has been given some attention recently. The procedures introduced by Bentler and Bonelt (1980) will be treated later in the study.

For external validity Burt (1973) did propose a statistical test which is based upon the structure involving only the observed variables. For structure involving both latent and observed variables, this statistic tends to "over-estimate" the predictive ability of the model. And it is impossible to differentiate the deviation of this over-estimation from the true one. Thus, the test for external validity was not performed in the study.
4.5 SUMMARY

In summary, this chapter described essentially the methodological considerations in the processes of building a model. These processes included specification, identification, parameters estimation, and testing of the model. The issues discussed in each section gave insights to the methods used in the study. The structural equations for the model were also specified in order to facilitate the understanding of the issues. This chapter has laid the groundwork for testing and modifications of the model, a topic to be treated in the next chapter.
Chapter V

MODEL TESTING

The previous chapter was devoted to the consideration of theories, methods, and data as related to the model in the study. This chapter will include testing the model and making refinements based on empirical data. The logic of model testing was treated in the latter part of Chapter IV. The procedures will be reviewed step by step in this chapter.

As shown earlier in Table 1, there were two measured exogenous variables, one measured endogenous variable, and five latent endogenous variables with two indicators for each. A total of thirteen measured indicators were used in the analysis. It was indicated that 119 T & I new and emerging occupations in manufacturing were identified. However, only ninety-eight were used in the study due to the availability of occupational statistics. The means and standard deviations of those measured indicators are shown in Table 2.

It was noted that much larger variation was found for those indicators with data being redistributed from industry to occupations. This could be explained, on the one hand, by the bi-directional variations (positive or negative) of some of the indicators and the sensitivity of small change rates of those indicators on the remaining ones.
<table>
<thead>
<tr>
<th>Indicator</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRDG</td>
<td>-0.0728</td>
<td>0.0794</td>
</tr>
<tr>
<td>CRDP</td>
<td>-0.0399</td>
<td>0.0560</td>
</tr>
<tr>
<td>CTFP</td>
<td>0.0187</td>
<td>0.0572</td>
</tr>
<tr>
<td>CPIR</td>
<td>0.0049</td>
<td>0.0179</td>
</tr>
<tr>
<td>CINO</td>
<td>0.0444</td>
<td>0.0947</td>
</tr>
<tr>
<td>CCLR</td>
<td>0.0297</td>
<td>0.0978</td>
</tr>
<tr>
<td>COSC</td>
<td>-1.4565</td>
<td>0.4736</td>
</tr>
<tr>
<td>COEC</td>
<td>-1.3250</td>
<td>0.3760</td>
</tr>
<tr>
<td>CWLE</td>
<td>2.3988</td>
<td>0.3117</td>
</tr>
<tr>
<td>CODE</td>
<td>-1.3328</td>
<td>0.6037</td>
</tr>
<tr>
<td>COSE</td>
<td>0.6263</td>
<td>0.2653</td>
</tr>
<tr>
<td>CELE</td>
<td>0.5495</td>
<td>0.1517</td>
</tr>
<tr>
<td>CNEO</td>
<td>0.0639</td>
<td>0.0190</td>
</tr>
</tbody>
</table>

The correlation coefficients matrix is shown in Table 3.

It was noted that the correlation coefficients between CNEO and all the indicators directly related to occupations (i.e., COSC, COEC, CODE, CWLE, CELE) were significant. The correlation coefficients between CNEO and other "indirect" indicators being redistributed from industry to occupations (except CCLR) were also significant at the 0.05 level but with much smaller magnitude. In addition, the correlation coefficients between indicators of a latent variable were also
TABLE 3
Correlation Coefficients of the Indicators

<table>
<thead>
<tr>
<th></th>
<th>CRDG</th>
<th>CRDP</th>
<th>CTFP</th>
<th>CPIR</th>
<th>CINO</th>
<th>CCLR</th>
<th>COSC</th>
<th>COEC</th>
<th>CWLE</th>
<th>CODE</th>
<th>COSE</th>
<th>CELE</th>
<th>CNEO</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRDG</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRDP</td>
<td></td>
<td>0.412*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTFP</td>
<td>-0.143</td>
<td>-0.329*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPIR</td>
<td>-0.339*</td>
<td>-0.153</td>
<td>0.094</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CINO</td>
<td>-0.194*</td>
<td>-0.131</td>
<td>0.905*</td>
<td>0.018</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCLR</td>
<td>-0.074</td>
<td>-0.051</td>
<td>0.834*</td>
<td>0.001</td>
<td>0.865*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COSC</td>
<td>0.310*</td>
<td>0.212*</td>
<td>-0.218*</td>
<td>-0.181</td>
<td>-0.262*</td>
<td>-0.182</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COEC</td>
<td>0.281*</td>
<td>0.338*</td>
<td>-0.316*</td>
<td>-0.160</td>
<td>-0.278*</td>
<td>-0.183</td>
<td>0.619*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CWLE</td>
<td>-0.320*</td>
<td>-0.217*</td>
<td>0.187*</td>
<td>0.154</td>
<td>0.180</td>
<td>0.140</td>
<td>-0.452*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CODE</td>
<td>0.196*</td>
<td>0.254*</td>
<td>-0.202*</td>
<td>-0.149</td>
<td>-0.144</td>
<td>-0.134</td>
<td>0.580*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COSE</td>
<td>-0.238*</td>
<td>-0.091*</td>
<td>0.160</td>
<td>-0.055</td>
<td>0.155</td>
<td>0.138</td>
<td>-0.328*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CELE</td>
<td>-0.229*</td>
<td>-0.248*</td>
<td>0.262*</td>
<td>-0.007</td>
<td>0.268*</td>
<td>0.173</td>
<td>-0.473*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNEO</td>
<td>-0.345*</td>
<td>-0.224*</td>
<td>0.240*</td>
<td>0.261*</td>
<td>0.270*</td>
<td>0.188</td>
<td>-0.676*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COEC</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CWLE</td>
<td>-0.521*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CODE</td>
<td>0.627*</td>
<td>-0.413*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COSE</td>
<td>-0.561*</td>
<td>0.581*</td>
<td>-0.488*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CELE</td>
<td>-0.684*</td>
<td>0.314*</td>
<td>-0.538*</td>
<td>0.466*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNEO</td>
<td>-0.751*</td>
<td>0.558*</td>
<td>-0.636*</td>
<td>0.544*</td>
<td>0.511*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at the 0.05 level

significant except for CTFP and CPTR. This gave a good indication that the model specified might not be far off the mark although it by no means guaranteed that the model was correctly specified.

In order to check and reinforce statistically (as contrasted to theoretically) the meaningful selection of variables in the model,
two pieces of information were needed. One was the proportion of variance of the dependent variables being explained by the independent variables. The other was the significant presence of individual variables. These pieces of information could be obtained by running a multiple regression with CNEO as the regressant and other indicators as the regressors. The results are shown in Table 4 and Table 5, respectively.

**TABLE 4**

Multiple Regression Using CNEO as Dependent Variable

<table>
<thead>
<tr>
<th>Sources</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>12</td>
<td>0.00206</td>
<td>17.30*</td>
</tr>
<tr>
<td>Residual</td>
<td>85</td>
<td>0.00012</td>
<td></td>
</tr>
</tbody>
</table>

$R = 0.842$  
$R^2 = 0.710$

*Significant at the 0.05 level.

It was shown in Table 4 that the multiple correlation ($R = 0.84$) was highly significant. The $R^2$ was 0.71, which meant that more than 70% of the variance of the change in distributions of T & I new and emerging occupations was explained by the independent variables. In terms of significant presence of the parameters, only the change rate of occupational structure (COSC), change rate of occupational employment change (COEC), change rate of demand for existing occupations (CODE), and change rate of patent intensity ratio (CPIR) were significant at the 0.05 level.
Based on the literature, it was not too surprising to see that COSC, COEC, and CODE achieved significance, but it was not expected for CPIR. However, the higher correlation coefficients of some pairs of independent variables as compared to those between the dependent and independent ones indicated the possibility of multicollinearity, which made the interpretation even more difficult.

The results were the empirical associations between the dependent and independent variables. These empirical associations were strictly reflections of the characteristics of the data in which the theories connecting the variables were not explicitly incorporated. Thus, using multiple regression in this case had little substantive

### TABLE 5

**Significant Presence of the Parameters**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>b</th>
<th>B</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>COES</td>
<td>-0.0200</td>
<td>-0.389</td>
<td>13.890*</td>
</tr>
<tr>
<td>COSC</td>
<td>0.0090</td>
<td>-0.228</td>
<td>7.289*</td>
</tr>
<tr>
<td>CODE</td>
<td>-0.0055</td>
<td>-0.174</td>
<td>4.214*</td>
</tr>
<tr>
<td>CPIR</td>
<td>0.1496</td>
<td>0.141</td>
<td>4.211*</td>
</tr>
<tr>
<td>COSE</td>
<td>0.0106</td>
<td>0.149</td>
<td>2.960</td>
</tr>
<tr>
<td>CINO</td>
<td>0.0595</td>
<td>0.297</td>
<td>2.532</td>
</tr>
<tr>
<td>CTFP</td>
<td>-0.0902</td>
<td>-0.272</td>
<td>2.071</td>
</tr>
<tr>
<td>CWLE</td>
<td>0.0051</td>
<td>0.084</td>
<td>1.105</td>
</tr>
<tr>
<td>CELE</td>
<td>-0.0080</td>
<td>-0.064</td>
<td>0.563</td>
</tr>
<tr>
<td>CRDG</td>
<td>-0.0034</td>
<td>-0.014</td>
<td>0.033</td>
</tr>
<tr>
<td>CRDP</td>
<td>-0.0023</td>
<td>-0.007</td>
<td>0.006</td>
</tr>
<tr>
<td>CCLR</td>
<td>-0.0003</td>
<td>-0.002</td>
<td>0.000</td>
</tr>
</tbody>
</table>

p(F > 3.97) = 0.05
implication for the problem at hand, especially when a priori knowledge
on the causal relationships of the variables was possibly postulated.
Nevertheless, the degree of empirical association derived from
multiple regression could at least indicate whether the model specifi-
cation was far off the empirical world, although it didn't tell how
close it was to reality.

5.1 PARAMETER ESTIMATES USING LISREL

The model testing, as indicated earlier, was essentially to test
how well the proposed model fit the empirical data. In other words,
it was to see whether the empirical behavior operated as suggested in
the model. In a multiple indicators model, the testing became one of
fitting the Σ imposed by the model to the sample matrix S with the
fitting function F to be minimized with respect to the vector of free
parameters. The fitting function F was described in equation (4.12)
in Chapter IV.

A computer program, LISREL, was used to estimate parameters and to
test and modify the model throughout the processes. LISREL is a general
computer program for estimating the unknown parameters in a set of
linear structural equations with latent and observed variables. In
addition, a chi-square statistic is provided for each model under
testing so that the goodness of fit of the model can be evaluated.

First of all, the model suggested in Figure 4 (model A1) and
formulated from equation (4.1) to (4.2) was tested. The model
conformed to all the basic assumptions of structural equation model,
namely, (a) uncorrelated disturbance terms, (b) uncorrelated latent variables and measured errors, and (c) uncorrelated measurement error terms. The results are shown in Table 6.

It was noted that the parameter values of one indicator of each latent variable were fixed as 1.0 for scaling purposes (Joreskog and Sorbom, 1978). The value of $p$ in Table 6 was interpreted as the probability of obtaining any chi-square value greater than the value actually obtained, given that the hypothesized model was true. In other words, the smaller the chi-square value, the higher the chances that the hypothesized model would be accepted. The perfect fit to the model occurs when the chi-square value is approaching zero on the one hand and the $p$ is approaching one on the other.

It was indicated in Table 6 that the hypothesized model based on Figure 4 was rejected since the chi-square value was too large to be acceptable. Thus, the parameter estimates in Table 6 didn't have any substantive meanings at this point. Nevertheless, they could serve as the sources for model modification, a step following immediately.

There was no standard procedure for multiple indicators model modification suggested in the literature, perhaps due to the relative newness of the approach. The method suggested by Sorbom (1975) in detecting correlated measurement errors looked promising for successive model modifications. This method is essentially to use the largest absolute value of the first-order derivative of the fitting function ($F$) to the elements of residual matrix of the measurement model as the source for modification. What it does is to treat the correlated measurement error having the largest absolute value as a free parameter.
TABLE 6  
Parameter Estimates for Model Al^1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>0.18</td>
<td>(0.15)²</td>
<td>b1</td>
<td>1.00</td>
<td></td>
<td>e1</td>
<td>-45.49</td>
<td>(-0.02)</td>
<td>r1</td>
<td>51.72</td>
<td>(0.02)</td>
</tr>
<tr>
<td>a2</td>
<td>0.31</td>
<td>(0.15)</td>
<td>b2</td>
<td>0.002</td>
<td>(0.02)</td>
<td>e2</td>
<td>1.00</td>
<td>(6.95)*</td>
<td>r2</td>
<td>1.18</td>
<td>(0.15)</td>
</tr>
<tr>
<td>a3</td>
<td>0.00</td>
<td>(0.02)</td>
<td>b3</td>
<td>1.00</td>
<td></td>
<td>e3</td>
<td>0.06</td>
<td>(3.03)*</td>
<td>r3</td>
<td>0.05</td>
<td>(4.03)*</td>
</tr>
<tr>
<td>a4</td>
<td>-0.06</td>
<td>(-0.13)</td>
<td>b4</td>
<td>0.92</td>
<td>(12.9)*</td>
<td>e4</td>
<td>0.20</td>
<td>(5.11)*</td>
<td>r4</td>
<td>0.25</td>
<td>(0.00)</td>
</tr>
<tr>
<td>a5</td>
<td>-3.52</td>
<td>(-0.03)</td>
<td>b5</td>
<td>1.00</td>
<td></td>
<td>e5</td>
<td>0.49</td>
<td>(6.11)*</td>
<td>r5</td>
<td>0.01</td>
<td>(0.00)</td>
</tr>
<tr>
<td>a6</td>
<td>-0.25</td>
<td>(-2.84)*</td>
<td>b6</td>
<td>1.16</td>
<td>(8.52)*</td>
<td>e6</td>
<td>0.19</td>
<td>(3.19)*</td>
<td>r6</td>
<td>-0.78</td>
<td>(-0.0)</td>
</tr>
<tr>
<td>a7</td>
<td>-40.2</td>
<td>(-0.00)</td>
<td>b7</td>
<td>1.00</td>
<td></td>
<td>e7</td>
<td>1.01</td>
<td>(6.96)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a8</td>
<td>71.3</td>
<td>(0.00)</td>
<td>b8</td>
<td>127.9</td>
<td>(0.05)</td>
<td>e8</td>
<td>0.63</td>
<td>(0.06)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a9</td>
<td>118.8</td>
<td>(0.05)</td>
<td>b9</td>
<td>1.00</td>
<td></td>
<td>e9</td>
<td>0.62</td>
<td>(5.83)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a10</td>
<td>54.4</td>
<td>(0.00)</td>
<td>b10</td>
<td>1.26</td>
<td>(5.32)*</td>
<td>e10</td>
<td>0.51</td>
<td>(4.20)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a11</td>
<td>602.</td>
<td>(0.001)</td>
<td>b11</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a12</td>
<td>100.</td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

x² = 179.06  df = 58  p = 0.00  N = 98

¹Model Al was the model proposed in Figure 4
²Values in parentheses are t statistics
*significant at 0.05 level
to be estimated in the next run. If a significant drop in chi-square is obtained, it indicates that the fit to the model is improved. The process continues until no significant chi-square drop is obtained.

The results of these successive model fittings are shown in Table 7.

It can be seen from Table 7 that the rule of stopping the fit at non-significant chi-squares was not followed exactly. However, the results indicated significant improvement in model fitting in terms of decreasing chi-square and increasing probability level (p) by freeing successively the parameters of correlated measured errors (from model A1 to A6).

Based on the results in Table 7, the model (A6) was still rejected at the 0.05 level. Thus, further refinement of the model seemed necessary. First, the measurement model was examined. Since the standardized values of the estimates in the $\lambda_y$ matrix are essentially the reliability coefficients of the indicators measured against their corresponding theoretical constructs (Wheaton et al, 1977), a close examination of these estimates would shed some light on possible sources for model revision. The standardized estimates of the matrix in nested models (A1 to A6) are shown in Table 8.

It was noted that almost all of the latent variables had reasonable reliabilities on their respective indicators except for the variable of technological change ($\lambda_1$ and $\lambda_2$). The reliability coefficient of CPIR ($\lambda_2$) was far off the reasonable range, indicating significant problem in using CPIR as the indicator. It seemed that
### TABLE 7

Model Fitting Allowing for Correlated Measurement Errors

<table>
<thead>
<tr>
<th>Model Mnemonic</th>
<th>Model</th>
<th>x^2</th>
<th>df</th>
<th>p</th>
<th>Decreased x^2</th>
<th>df^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0</td>
<td>mutual independence</td>
<td>597.51</td>
<td>70</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>uncorrelated error</td>
<td>179.06</td>
<td>57</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>A1, ( \psi_{73} )</td>
<td>109.67</td>
<td>56</td>
<td>0.000</td>
<td>69.79*</td>
<td>1.9584</td>
</tr>
<tr>
<td>A3</td>
<td>A2, ( \psi_{93} )</td>
<td>99.96</td>
<td>55</td>
<td>0.0002</td>
<td>9.71*</td>
<td>1.8175</td>
</tr>
<tr>
<td>A4</td>
<td>A3, ( \psi_{61} )</td>
<td>97.87</td>
<td>54</td>
<td>0.0002</td>
<td>2.09</td>
<td>1.8124</td>
</tr>
<tr>
<td>A5</td>
<td>A4, ( \psi_{106} )</td>
<td>87.36</td>
<td>53</td>
<td>0.0021</td>
<td>10.51*</td>
<td>1.6483</td>
</tr>
<tr>
<td>A6</td>
<td>A5, ( \psi_{98} )</td>
<td>84.52</td>
<td>52</td>
<td>0.0045</td>
<td>2.84</td>
<td>1.5869</td>
</tr>
<tr>
<td>B1^2</td>
<td>A6, ( \psi_{2} )</td>
<td>46.58</td>
<td>41</td>
<td>0.2537</td>
<td>37.94*</td>
<td>1.1361</td>
</tr>
</tbody>
</table>

1 The correlated error indicated was relaxed as free parameter to be estimated
2 The indicator of CPIR (\( \psi_{2} \)) was excluded from the model
3 Significant at the 0.05 level

### TABLE 8

Reliability Coefficients in Successive Measurement Models

<table>
<thead>
<tr>
<th>Model</th>
<th>( \lambda_1 )</th>
<th>( \lambda_2 )</th>
<th>( \lambda_3 )</th>
<th>( \lambda_4 )</th>
<th>( \lambda_5 )</th>
<th>( \lambda_6 )</th>
<th>( \lambda_7 )</th>
<th>( \lambda_8 )</th>
<th>( \lambda_9 )</th>
<th>( \lambda_{10} )</th>
<th>( \lambda_{11} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>0.97</td>
<td>0.06</td>
<td>0.96</td>
<td>0.90</td>
<td>0.71</td>
<td>0.88</td>
<td>0.73</td>
<td>0.61</td>
<td>0.65</td>
<td>0.74</td>
<td>1.00</td>
</tr>
<tr>
<td>A2</td>
<td>2.69</td>
<td>0.03</td>
<td>0.96</td>
<td>0.89</td>
<td>0.74</td>
<td>0.88</td>
<td>0.75</td>
<td>0.61</td>
<td>0.57</td>
<td>0.70</td>
<td>1.00</td>
</tr>
<tr>
<td>A3</td>
<td>4.25</td>
<td>0.02</td>
<td>0.96</td>
<td>0.89</td>
<td>0.71</td>
<td>0.87</td>
<td>0.75</td>
<td>0.61</td>
<td>0.66</td>
<td>0.73</td>
<td>1.00</td>
</tr>
<tr>
<td>A4</td>
<td>3.52</td>
<td>0.03</td>
<td>0.96</td>
<td>0.89</td>
<td>0.72</td>
<td>0.87</td>
<td>0.74</td>
<td>0.61</td>
<td>0.68</td>
<td>0.73</td>
<td>1.00</td>
</tr>
<tr>
<td>A5</td>
<td>3.12</td>
<td>0.04</td>
<td>0.96</td>
<td>0.89</td>
<td>0.72</td>
<td>0.82</td>
<td>0.70</td>
<td>0.59</td>
<td>0.68</td>
<td>0.62</td>
<td>1.00</td>
</tr>
<tr>
<td>A6</td>
<td>0.95</td>
<td>0.02</td>
<td>0.96</td>
<td>0.90</td>
<td>0.74</td>
<td>0.84</td>
<td>0.73</td>
<td>0.57</td>
<td>0.68</td>
<td>0.70</td>
<td>1.00</td>
</tr>
</tbody>
</table>
there was a connection between, on the one hand, this extreme
"unreliability" and the significant contribution of CPIR to CNEO in
earlier regression analysis and nonsignificant correlation coefficient
between CPIR and CTFP on the other. However, the mathematical
exposition for this phenomenon was yet to be determined. Also, the
t values of the estimates of this indicator in various models were
not significant from zero. In the meantime, the reliability coefficients
of CTFP (λ₂) were fluctuating from one model to the others. This might
be the effect of extremely small reliability coefficients of CPIR,
since both indicators represented the same latent variable.

Considering the nonsignificant reliability coefficient and
possible "spill-over" effect of CPIR on the other indicator, it was
decided that the indicator of CPIR should be discarded. Only the CTFP
was used to represent the variable of technological change. This
decision was purely an empirical one since patent intensity was
suggested by Mansfield (1968) as a possible measure for technological
change.

The model fitting statistics are also shown in Table 7 under the
B1 model (excluding CPIR). It was noted that the chi-square dropped
significantly (39.94) while omitting the indicator CPIR. The probability
level of this model was 0.2537, which indicated that the model under
B1 specification had a good fit to the data. The successive model
fitting processes were repeated for models without CPIR and similar
results were obtained. This confirmed the inclusions of correlated
measurement errors in the model. The final estimates of the parameters
and their t values are shown in Table 9. The interpretations of these parameters were treated in the next section.

Since the goodness of fit of nested models is based on large sample theory (the chi-square variate is a direct function of sample size), the probability of rejecting any model increases as the sample size increases. For a model with a moderately small sample, the inference is different. Therefore, additional information was needed to validate the goodness of fit when the sample size was not sufficiently large. Bentler and Bonett (1980) proposed an index called Incremental Fit Index, which provided the information about the amount of information gained in successive model fitting. This index is independent of sample size and can reflect the goodness of fit of competing models. The formula for calculating this index is as follows:

\[ \rho_{kl} = \frac{Q_k - Q_0}{Q_0 - 1} \]  \hspace{1cm} (5.1)

where \( \rho_{kl} \) : incremental fit index of model k over model 1

\( Q_k \) : \( X^2/df \) ratio of model k

\( Q_0 \) : \( X^2/df \) ratio of model 1

It was noted from equation (5.1) that the magnitude of \( \rho_{kl} \) is heavily dependent on the magnitude of \( Q_0 \). Thus, the specification of the null model becomes critical. It was suggested that the null model is the model with mutual independence of the latent variables (Bentler and Bonett, 1980). In other words, the parameters in the B
Table 9
Parameter Estimates for Model Bl

<table>
<thead>
<tr>
<th></th>
<th>$a_i$</th>
<th>$b_i$</th>
<th>$\lambda_i$</th>
<th>$e_{ij}$</th>
<th>$r_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>0.065 (1.94)*</td>
<td>1.000</td>
<td>0.944</td>
<td>0.111 (2.58)*</td>
<td>0.046 (-1.22)</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.256 (6.82)*</td>
<td>1.000</td>
<td>0.955</td>
<td>0.071 (3.31)*</td>
<td>0.532 (1.50)</td>
</tr>
<tr>
<td>$a_3$</td>
<td>0.831 (2.39)*</td>
<td>0.938 (17.01)*</td>
<td>0.896</td>
<td>0.198 (6.11)*</td>
<td>0.491 (4.04)*</td>
</tr>
<tr>
<td>$a_4$</td>
<td>-0.241 (-0.96)</td>
<td>1.000</td>
<td>0.740</td>
<td>0.453 (6.14)*</td>
<td>-0.010 (-0.10)</td>
</tr>
<tr>
<td>$a_5$</td>
<td>-0.980 (-19.8)*</td>
<td>1.134 (8.46)*</td>
<td>0.839</td>
<td>0.298 (4.84)*</td>
<td>0.006 (0.09)</td>
</tr>
<tr>
<td>$a_6$</td>
<td>-0.585 (-16.8)</td>
<td>1.000</td>
<td>0.729</td>
<td>0.505 (4.89)*</td>
<td>0.223 (2.54)*</td>
</tr>
<tr>
<td>$a_7$</td>
<td>-0.089 (-0.09)</td>
<td>-0.778 (-5.67)*</td>
<td>-.567</td>
<td>0.678 (6.42)*</td>
<td></td>
</tr>
<tr>
<td>$a_8$</td>
<td>0.909 (1.45)</td>
<td>1.000</td>
<td>0.680</td>
<td>0.556 (5.34)*</td>
<td></td>
</tr>
<tr>
<td>$a_9$</td>
<td>0.773 (5.85)*</td>
<td>1.035 (5.58)*</td>
<td>0.704</td>
<td>0.497 (4.99)*</td>
<td></td>
</tr>
<tr>
<td>$a_{10}$</td>
<td>1.185 (1.00)</td>
<td>1.000</td>
<td>1.00</td>
<td>-0.030 (-1.55)</td>
<td></td>
</tr>
<tr>
<td>$a_{11}$</td>
<td>0.398 (0.74)</td>
<td></td>
<td></td>
<td>0.074 (2.75)*</td>
<td></td>
</tr>
<tr>
<td>$a_{12}$</td>
<td>0.483 (0.75)</td>
<td></td>
<td></td>
<td>-0.037 (-1.39)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.206 (2.62)*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.150 (-2.58)*</td>
<td></td>
</tr>
</tbody>
</table>

$x^2 = 46.58$  \hspace{1cm} df = 41  \hspace{1cm} p = 0.2537  \hspace{1cm} N = 98

Values in parentheses are t statistics
*Significant at 0.05 level
and \( \Gamma \) matrix were to be fixed to zeros. The results of these incremental fit indices with respect to different model comparisons are shown in Table 10.

**TABLE 10**

Incremental Fit Indices of Model Comparisons

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>( \Delta \chi^2 )</th>
<th>df</th>
<th>( \rho_{kl} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0 - A1</td>
<td>377.67</td>
<td>13</td>
<td>0.716</td>
</tr>
<tr>
<td>A0 - A2</td>
<td>447.06</td>
<td>14</td>
<td>0.873</td>
</tr>
<tr>
<td>A0 - A3</td>
<td>456.77</td>
<td>15</td>
<td>0.891</td>
</tr>
<tr>
<td>A0 - A4</td>
<td>458.86</td>
<td>16</td>
<td>0.892</td>
</tr>
<tr>
<td>A0 - A5</td>
<td>469.37</td>
<td>17</td>
<td>0.914</td>
</tr>
<tr>
<td>A0 - A6</td>
<td>474.21</td>
<td>18</td>
<td>0.922</td>
</tr>
<tr>
<td>A0 - B1</td>
<td>550.93</td>
<td>29</td>
<td>0.982</td>
</tr>
</tbody>
</table>

It was noted that the improvements in fit obtained by various models over A0 (the null model) ran from 0.716 for A1 to 0.982 for B1. It was a significant improvement for B1 since the remaining improvement (1-0.982 = 0.018) that might be obtained from a more adequate model was insignificant from a practical point of view. Thus, the model B1 was both statistically as well as practically accepted.

5.2 **INTERPRETATIONS OF THE RESULTS**

This section dealt with the interpretations of the model based on the parameters estimated in the preceding section. These interpretations followed closely the hypotheses proposed at the end of Chapter II.
Overall, the model Bl fits the data quite well from both statistical and practical viewpoints. In other words, the probability of obtaining an alternative model which might fit the data even better was nonsignificant on the one hand and the room for improvement of fit was very small on the other. This result could not be overstated because the sample size was moderately small and the sample was biased due to the availability of occupational statistics. Therefore, the estimates were tentative and the results should be cautiously interpreted.

First, it began with the interpretation of the measurement model. It was noted that all of the measurement parameters were significantly different from zero (see Table 9). The ranges of value for these measurement parameters looked reasonable except for CODE having a negative sign. It indicated that the average change rate of demand of existing occupations was a cause indicator rather than an effect one. It was contrary to the proposition that the theoretical construct of demand effects had two dimensions of effect, one the change of demand while the other the change of wages for those existing occupations. It seemed, according to the data, that the change of demand of existing occupations was a cause for any demand effect to be taking place. This certainly would lead to a chicken-and-egg type of argument. It also displayed the fundamental difficulty of using such a method while firm theory concerning causal priorities was not a priori established. Since there was no convincing evidence to select one over the other, plus the consideration of the effect of a biased sample, the result was rather inconclusive at this point.
In terms of the reliabilities of other indicators, it was shown in Table 9 that most of them fell into reasonable range of value. However, there were differences between indicators measuring the same theoretical construct. The pair of indicators measuring sectoral growth (CINO, CCIR) performed very well in terms of their reliability coefficients (0.955, 0.896, respectively). For the variable of occupational structure change, the indicator of COEC performed a little better than it did for COSC (0.839 compared to 0.740). However, the difference was not too large. For the variable of demand effect of existing occupations, there was a large discrepancy between the reliabilities of CWLE and CODE (0.729 and 0.567 in absolute value). This might be due to the fact that the CODE was specified an effect indicator, but the data indicated otherwise. For the variable of supply effect of existing occupations, it was surprising to see that the indicators of COSE and CELE performed quite well (0.680 and 0.704, respectively) considering the fact that the inclusion of them was exploratory in nature. So, in general, the measurement model performed reasonably well in terms of the reliability coefficients.

Next, the structure model should be examined. This examination followed exactly the hypotheses proposed. The parameter estimates shown in Table 9 (a_i) for structural parameters could be interpreted similarly to those partial regression coefficients in ordinary least square analysis. They could be interpreted as one unit change of the independent variable resulting in the amount of unit change of dependent variable indicated by the parameter estimate.
First, it was surprising to discover that the first three hypotheses proposed in Chapter II failed to achieve significance. In other words, the demand effects, the supply effects, and the occupational structure change of existing occupations didn't have significant direct effects \((a_{11}, a_{12}, \text{ and } a_8, \text{ respectively})\) on the change in distributions of T & I new and emerging occupations in manufacturing. In other words, the labor market process did not have significant impact on the development of T & I new and emerging occupations. Also, the impact of technological change did not get propagated to the development of T & I new and emerging occupations. This was contrary to what the theories and evidence suggested. An examination of the magnitudes of their t statistics revealed that the occupational structure change of existing occupations was close to achieving significance \((t = 1.45)\). It was extremely difficult to assess what caused the discrepancy between the theories and the data. Either weak theories or poor quality of data or both could be responsible for this result. It seemed logical at this point to suspect the data first, since the sample was biased and the sample size was moderately small. This led to a larger standard of error of estimate and, consequently, a smaller t value. Nevertheless, this surfacing of discrepancy between theories and data suggested that a closer examination of the theories which led to the formulation of the model and a better set of data were needed.

The fourth hypothesis that the occupational structure change had significant indirect effect on the development of T & I new and emerging occupations through the variables of demand and supply
effects was not accepted. The data indicated an opposite causation (the arrow might run from demand effect of existing occupations to occupational structure change instead of the one hypothesized since \( a_7 \) was negative) despite the nonsignificance of the parameter. This at least suggested that occupational structure change might have only a direct effect on the development of T & I new and emerging occupations.

The fifth hypothesis that significant reciprocal causal relationships between demand and supply effects was half-accepted. There was a significant causation from demand effects to the supply effects (\( a_9 \) was significant). However, the parameter (\( a_{10} \)) of supply effects impacting on the demand effects was not significant. The former part confirmed the substitution effect (the more the demand, the higher the wage, the more the labor supply) in the traditional labor economics framework. However, the later part that more labor supply would burst more labor demand due to the inducement of substitution of labor for capital was not sustained here. It was conceivable that this might be true only if the elasticity of substitution between labor and capital was high. Since there was no evidence to indicate preference of one way over the other, the interpretation for this nonsignificance was at least tentatively acceptable.

The sixth hypothesis dealt with the significant direct effect of sectoral growth on occupational structure change. It was hypothesized that growth in industries had direct impact on change in occupational structure. However, the data (\( a_6 \)) showed the reverse: that changes in occupational structure had a direct impact on sectoral
growth although its influence failed to achieve significance. This result implied an interesting inference that changes in occupational structure due to technological change might be a force actually stimulating the growth in industries rather than a result of it. This point needed further investigation, however.

The seventh hypothesis related to significant reciprocal causal relationships between technological change and sectoral growth was also half-accepted. Since both the parameters \(a_4\) and \(a_5\) were negative as shown in Table 9, the causal interpretations were made in accordance with the direction of influence. It was noted that technological change did have a significant impact on sectoral growth as hypothesized. However, the reciprocal relationship that sectoral growth had significant impact on technological change was not established. This nonsignificant impact of sectoral growth on technological change was different from what Kendrick (1973) suggested. However, the context and method of this study were different from Kendrick's one and a direct comparison might not be appropriate.

The eighth hypothesis that technological change had a significant direct impact on occupational structure change was accepted. This was evidenced by the significant parameter estimate \(a_3\) shown in Table 9. In other words, technological change did have significant impact on changes in occupational structure according mainly to the job generating process as outlined in Chapter II.

The last hypothesis dealt with the significant contribution of government R & D and private company R & D to technological change. It was indicated that the private company R & D \(a_2\) had a more
significant impact on technological change than it did for government
R & D ($a_1$) since the latter barely achieved significance. This
result was consistent with the conclusions made by Terleckyj (1980)
although slightly different measures on the variables were used.
In addition, there was a significant correlation between government
and private company's changes in R & D expenditure intensity ratios
(see Table 2). This could be explained by noting that the level of
investment in R & D by private companies was highly correlated with
the level of investment in R & D by the government during the period
studied.

The results made in this section could not be claimed without
further qualification in terms of the effect of correlated measurement
errors. Essentially the fit of the model was obtained at the expense
of allowing for spurious correlations among errors of indicators. It
could be seen from Table 9 that five correlated measurement errors
were relaxed from zero ($e_{61}, e_{73}, e_{93}, e_{98}, e_{106}$). And three of them
were statistically different from zero ($e_{73}, e_{98}, e_{106}$).

The interpretations of these correlated measurement errors were
very difficult to make since the source of nonrandom measurement
ersors were not known. For example, it was harder to explain why there
existed a significant correlation ($e_{73}$) between the error terms of
change rates of industry output and wage level of existing occupations
than there did ($e_{98}$) for the error terms of change rates of supply and
demand of existing occupations. For the latter, there were variables,
(e.g., wage, utilities of work and leisure, etc.) intermingling the
labor supply and demand interaction. In other words, the latter
correlated measurement error would be anticipated a priori but not for the former. Similarly, it was difficult to explain the correlated measurement error \( e_{106} \) between the change rates of educational level and occupational employment change of the existing occupations.

Costner and Schoenberg (1973) indicated that it was impossible to detect the source for correlated measurement errors in a multiple indicators model using only two indicators for each latent variable.

It was found in the literature review that most authors avoided interpreting in detail the correlated measurement errors. It seemed that if there was no particular reason to specify a priori any correlated measurement errors, relaxing these parameters was merely a means to bring the chi-square down to an acceptable level on the one hand and to detect correlated measurement errors on the other. This ad hoc practice was understandable, since in most cases, the interpretations of correlated measurement errors was really obscure. Therefore, no attempt was made here to interpret further these potentially troublesome correlated errors.

5.3 SUMMARY

In summary, this chapter dealt with data analysis in the context of model testing. The descriptive statistics in terms of means, standard deviations, and correlational coefficients among indicators were first presented. A multiple regression was run next to establish the empirical associations in terms of variance explanation of the dependent variables. Then, the LISREL program was used to estimate
the parameters and to make revisions on the model. The goodness of
fit of the model and the interpretations of the parameters estimated
in light of the hypotheses were treated last.

The goodness of fit of the model based on statistical (chi-
square) and practical (Incremental Fit Index) criteria was well
accepted. The reliability coefficients of the indicators representing
the latent variables were in reasonable range. The parameter estimates
representing the major hypotheses of impact of occupational structure
change, demand effects, and supply effects on changes in distributions
of T & I new and emerging occupations in manufacturing were not
statistically significant. However, their magnitudes were not too far
from achieving significance.

In general, the causal development of T & I new and emerging
occupations in manufacturing followed more closely to the economic
force of technological change than it did to labor market process
of demand and supply interactions. However, owing to the problems
in data and some nonrandom measurement errors in the model, the
results should be considered directive instead of conclusive.
Chapter VI

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

6.1 SUMMARY

The purpose of this study was to develop a model which could shed light on the causal relationships among the factors impacting on the development of T & I new and emerging occupations in manufacturing. The rationale was to accumulate knowledge for better understanding of the process by which new and emerging occupations have been developed or have emerged. Once the contributions of the factors impacting on the development have been established, their progress can be monitored. Consequently, a forecast with respect to the quantity of distribution of new and emerging occupations that would develop or emerge at a certain time in the future could be made.

Since the development of any causal model is a result of interweaving theories, methods, and data, the literature review was focused on these three areas. The variables included in the model were selected based on the theories, empirical evidence, and postulations gathered from literature review. There were two economic forces, technological change and labor market process, by which the development of new and emerging occupations was suggested. The labor market process corresponded to the traditional labor demand and supply
framework in which the demand and supply effects of occupations being substitutable by the new and emerging ones were the major forces. The technological change took the shift in employment structure and the job generation process in which the occupational structure change, was directly impacting on the development of new and emerging occupations. In addition, technological change was not exogenous but rather endogenous to the process. A host of exogenous factors were impacting on technological change. However, only R & D in government and private companies were included in the model due to their consistent significant impact gathered from various studies and to the compactness of the model. A conceptual model was developed based on these two economic forces. Hypotheses were formulated based on the directions of impact between variables.

During the process of determining variables in the model, it was found that the majority of the variables were not directly observable (latent variables). Thus, multiple indicators representing each latent variable were necessary. This led to building a multiple indicators causal model based on the conceptual one. There were five latent endogenous variables with ten indicators, one measured endogenous variables, and two measured exogenous variables.

The study was exploratory in nature since little empirical evidence was obtained beforehand concerning the statistical associations among those variables. Thus, it was decided that the data analysis would be in two stages. First, a simple multiple regression was done, using the change in distributions of T & I new and emerging occupations
as the dependent variable to regress on the independent variables, which happened to be the indicator selected in the model. This would give indications concerning the empirical associations of those indicators. Second, the parameters in the multiple indicators model were estimated using the LISREL program, the most appropriate canned program in dealing with a model with latent and measurement structures. In the meantime, the fit of the proposed model to data was examined and revision of the model was made based on chosen criteria.

The data were obtained from different government sources and from individual research data sets. Data transformations or adjustments were needed on occasions where the original forms of the data were not directly applicable. The details of transformation were outlined in Chapter III. The population of this study were the T & I new occupations which appeared in the fourth edition of the Dictionary of Occupational Titles, but not in the third edition. This determined the time frame of the study being from 1965 to 1977, since these were the years that respective editions of the DOT were published. The sample was not randomly drawn due to limited occupational statistics available to the population (data from Occupational Employment Statistics Program were used). This was the problem that potentially biased the parameter estimates. However, it was a necessary tradeoff in order to pursue the study.

6.2 CONCLUSIONS

The examination of the model explaining the development of T & I new and emerging occupations in manufacturing during the period of
concern was made from three levels. The first level was the overall
goodness of fit of the proposed model to the data. The second level
was the adequacy of the measurement structure. And the third level
was the appropriateness of the latent structure.

The goodness of fit of the proposed model to the data was
judged against both the statistical and practical criteria. The
statistical criterion was the chi-square statistics and its associated
probability level, while the practical criterion was the Incremental
Fit Index. It was shown in Table 7 that model B1 had a decent chi-square
(46.58) given the associated degrees of freedom (41). And the
probability of obtaining an alternative model which might fit the data
even better was nonsignificant. Also shown in Table 10 was a high
Incremental Fit Index (0.982), which indicated very little room for
further model modification. Thus, the goodness of fit of model B1 was
well accepted.

The adequacy of the measurement structure of the model was judged
by the reliability coefficients of the indicators with respect to their
latent variables. The change rate of patent intensity ratio, one of
the indicators for technological change, was discarded due to its
extremely low reliability coefficient. It was shown in Table 9 that
all of the indicators were significantly different from zeros and in
reasonable range of value. Only the change rate of demand of existing
occupations had a negative value. This indicated a difference between
the theory and data since the indicator was proposed as an effect
indicator (the negative value meant a cause indicator). Conceptually
speaking, an argument could be made whether this change rate of demand
of existing occupations was an effect indicator of the demand effects or a cause of it. Thus, the conclusion related to this indicator was made tentative, pending further evidence. Nonetheless, the remaining ones were reasonably reliable indicators for respective latent variables.

The most important examination of the model was the appropriateness of the latent structure. This linked directly to the hypotheses proposed for possible explanation of the development of T & I new and emerging occupations in manufacturing. The appropriateness of this latent structure was judged by the proposed direction of influence of the parameters being estimated and their significance. Unfortunately, the three most important hypotheses concerning the direct impact on the development of T & I new and emerging occupations were not statistically accepted. The occupational structure change and the demand and supply effects did not impact significantly on the development of T & I new and emerging occupations although the directions of causation were correctly specified. In other words, the labor market process did not have significant impact on the development of T & I new and emerging occupations. Also the impact of technological change did not get propagated significantly, via the occupational structure change, to the development of new and emerging occupations. This was contrary to what the theories suggested in Chapter II. This nonsignificance should be examined with two possibilities. One was poor formulation of theory while the other was poor quality of data. Since the directions of causation were correctly specified and the parameters were not too far from achieving
statistical significance, one would tend to suspect the data. In fact, there were potential problems associated with the quality of data since the sample size was moderately small, the sample was biased, and a lot of data transformations were made before model fitting. This potentially poor quality of data would result in large variance of the parameter estimates and, consequently, more conservative t values were obtained. Therefore, further verification of the parameters estimated was definitely needed.

The hypothesis concerning the reciprocal relationships between demand and supply effects was half-accepted. Only the significant impact from demand to supply was established. This was consistent with the substitution effect in labor economics. The proposition of supply causing demand was always contingent upon certain conditions to be satisfied (in this case, a high elasticity of substitution between labor in existing occupations and capital). Since there was no evidence indicating that this condition was satisfied, the non-significance of supply effects on demand effects was not too surprising.

The hypothesis that occupational structure change had indirect effect on development of T & I new and emerging occupations through the variables of demand and supply effects was not supported by the data. This indicated that no strong interaction existed between technological change and labor market process. Additionally, it was not possible to accept the hypothesis of direct impact of sectoral growth on occupational structure change. This indicated that change in industrial output did not impact significantly on shifting distributions of occupations in industries and changing employment patterns in the economy.
The hypothesis that technological change had a significant impact on sectoral growth was accepted. This was consistent with the theory that technological change led to changes in composition of the industries. However, the opposite was not accepted. This result did not support the notion that more sectoral growth would lead to more resources for investment, and subsequently more technological change.

The hypothesis that technological change had significant impact on occupational structure change was accepted. This indicated that technological change had significant contribution to shifting mix of occupations in industries and changing employment patterns in the economy. Also accepted was the hypothesis that R & D expenditure in private companies and government had significant impact on technological change. And the former was more significant than the latter. This was consistent with the empirical results suggested by other researchers. In addition, a significant correlation between R & D expenditures in private companies and government was established.

In summary, it seemed, based on the data, that the causal development of T & I new and emerging occupations in manufacturing followed more closely to the economic force of technological change than it did to the labor market process of demand and supply interactions. There were policy implications derived from these conclusions for vocational educators and planners. First, it is recognized that the linkage between vocational educators and labor market process, especially the supply side, is tighter than it is between vocational educators and technological change. In other words, vocational educators have more influence on the development of T & I new and
emerging occupations by supplying qualified workers for substitutable occupations than they do by influencing technological change. However, there was failure in this study to establish a significant impact of supply effects on development of T & I new and emerging occupations. This meant that vocational educators could do very little in facilitating the developmental behavior. However, this did not mean that vocational educators could escape from the responsibility of preparing qualified workers for potential new and emerging occupations. Second, vocational educators could at least influence indirectly the developmental behavior by advocating policy; for instance, tax credit for R & D investment, which could facilitate technological change. This was due to the fact that technological change would more likely have greater impact on the development of T & I new and emerging occupations in manufacturing.

A word of caution should be given in relationship to the results of this study. Owing to the data problem indicated earlier, several significant nonrandom measurement errors in the model, and the unstableness of maximum likelihood estimates from alternative specifications of the model, the results should be considered rather tentative at this point. Until a population associated with better occupational statistics from which a random sample can be obtained, the parameter estimates should only be regarded as directive instead of conclusive.
6.3 RECOMMENDATIONS

The recommendations based on the results of this study were as follows:

1. It was recommended that effort from the Bureau of Labor Statistics to collect occupational statistics in OES occupational codes with built-in provisions for new and emerging occupations be continued. The most severe problem encountered in the study was the lack of a quality data set in occupational statistics corresponding to the new and emerging occupations (whatever defined). The crosswalk between occupational classification systems was at best an approximation. Since the development of new and emerging occupations in the labor market is an evolving process, it would be ideal if the occupational employment statistics would have been collected periodically using the appropriate occupational classification. Since on the one hand, the Census occupational codes are too generic to be useful in this regard and the DOT occupational codes are too detailed to collect on the other, it seemed that the OES occupational codes were the most appropriate for a study of this nature.

2. It was recommended that a continuous effort to refine the theories and to obtain empirical evidence related to the structure of the latent variables and the structure between latent variables and their indicators be pursued. It seemed that decomposing the model into several blocks and then obtaining empirical evidence to support the causal orderings of the variables in each block would be a
fundamental way to validate the specification of the latent structure and the measurement structure of the model.

3. It was recommended that similar studies be extended to new and emerging occupations in other service areas in vocational education and in different industries or economic sectors so that a comparison can be made in terms of the significant impact of the causal factors.

4. It was recommended that similar studies be conducted using regional or state data. This would establish knowledge concerning the patterns of development of new and emerging occupations that relates particularly to the region or the state.

5. This model could at best be used to project the quantity (proportion of the industry employment) of T & I new and emerging occupations in manufacturing, given the validity of the model to be established. It was recommended that a way to obtain, out of the projected quantity, the quality information (types of new and emerging occupations being distributed) be pursued so that a base for curriculum development can be established.

6. It was recommended, given that the validity of the model is established, that changes in the causal factors proposed in the model be monitored by vocational educators. For example, the data for technological change in terms of the changes in Total Factor Productivity in different industries and the change of occupational structure in terms of changes in occupational structure shift and occupational employment shift of those existing (and substitutable) occupations
should be collected and examined constantly. This could raise vocational educators' awareness of potential changes in development of new and emerging occupations. Consequently, a vocational delivery system more responsive to the labor market signals would be in place.
APPENDIX

Trade and Industrial New and Emerging Occupations in Manufacturing
## Trade and Industrial New and Emerging Occupations in Manufacturing

<table>
<thead>
<tr>
<th>NO.</th>
<th>DOT</th>
<th>DOT TITLE</th>
<th>OES SVP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>500380010</td>
<td>Plater</td>
<td>55C29 7</td>
</tr>
<tr>
<td>2</td>
<td>504682010</td>
<td>Annealer</td>
<td>55J85 4</td>
</tr>
<tr>
<td>3</td>
<td>510465010</td>
<td>Carbide-Powder Processor</td>
<td>55S04 5</td>
</tr>
<tr>
<td>4</td>
<td>512467010</td>
<td>Potline Monitor</td>
<td>55M43 4</td>
</tr>
<tr>
<td>5</td>
<td>513362010</td>
<td>Calciner Operator</td>
<td>55P40 6</td>
</tr>
<tr>
<td>6</td>
<td>5143600010</td>
<td>Die-Casting-Machine Setter</td>
<td>55R73 7</td>
</tr>
<tr>
<td>7</td>
<td>520382014</td>
<td>Liquid-Sugar Melter</td>
<td>55B93 4</td>
</tr>
<tr>
<td>8</td>
<td>520585022</td>
<td>Liquid-Sugar Fortifier</td>
<td>55B93 4</td>
</tr>
<tr>
<td>9</td>
<td>520685074</td>
<td>Cocoa-Powder-Mixer Operator</td>
<td>55B93 4</td>
</tr>
<tr>
<td>10</td>
<td>521362014</td>
<td>Miller, Distillery</td>
<td>55L30 4</td>
</tr>
<tr>
<td>11</td>
<td>521382010</td>
<td>Evaporator Operator</td>
<td>55I39 5</td>
</tr>
<tr>
<td>12</td>
<td>522382038</td>
<td>Vinegar Maker</td>
<td>55B93 4</td>
</tr>
<tr>
<td>13</td>
<td>522662010</td>
<td>Receiver, Fermenting Cellars</td>
<td>55S43 4</td>
</tr>
<tr>
<td>14</td>
<td>522685062</td>
<td>Fermenter</td>
<td>55S43 4</td>
</tr>
<tr>
<td>15</td>
<td>522685098</td>
<td>Still Operator</td>
<td>55Q11 4</td>
</tr>
<tr>
<td>16</td>
<td>523682022</td>
<td>Drier Operator</td>
<td>55D91 4</td>
</tr>
<tr>
<td>17</td>
<td>534682018</td>
<td>Coating-Machine Operator</td>
<td>55G43 4</td>
</tr>
<tr>
<td>18</td>
<td>539367014</td>
<td>Water-Quality Tester</td>
<td>55Q57 5</td>
</tr>
<tr>
<td>19</td>
<td>552665014</td>
<td>Batch-Still Operator 1</td>
<td>55Q11 4</td>
</tr>
<tr>
<td>20</td>
<td>553364010</td>
<td>Sample Tester</td>
<td>55Q57 5</td>
</tr>
<tr>
<td>21</td>
<td>558685026</td>
<td>De-Ionized Operator</td>
<td>55S76 4</td>
</tr>
<tr>
<td>22</td>
<td>558685038</td>
<td>Ion-Exchange Operator</td>
<td>55G15 5</td>
</tr>
<tr>
<td>23</td>
<td>559682010</td>
<td>Capsule-Filling-Machine Operator</td>
<td>55C39 4</td>
</tr>
<tr>
<td>24</td>
<td>562685018</td>
<td>Operator, Prefinish</td>
<td>55R68 5</td>
</tr>
<tr>
<td>25</td>
<td>570682014</td>
<td>Plant Operator</td>
<td>55A67 4</td>
</tr>
<tr>
<td>26</td>
<td>673382014</td>
<td>Spray-Drier Operator</td>
<td>55S38 4</td>
</tr>
<tr>
<td>27</td>
<td>575365010</td>
<td>Glass-Ribbon-Machine Operator</td>
<td>55I19 6</td>
</tr>
<tr>
<td>28</td>
<td>575382018</td>
<td>Glass-Bulb-Machine Former</td>
<td>55I19 5</td>
</tr>
<tr>
<td>29</td>
<td>579384018</td>
<td>Ware Tester</td>
<td>55Q57 5</td>
</tr>
<tr>
<td>30</td>
<td>582482010</td>
<td>Colorer, Hides and Skins</td>
<td>55Q50 4</td>
</tr>
<tr>
<td>31</td>
<td>582482014</td>
<td>Tanner, Rotary Drum</td>
<td>55K74 5</td>
</tr>
<tr>
<td>32</td>
<td>582685158</td>
<td>Warp-Dying-Vat Tender</td>
<td>55I13 4</td>
</tr>
<tr>
<td>33</td>
<td>590684018</td>
<td>Etched-Circuit Processor</td>
<td>55I35 6</td>
</tr>
<tr>
<td>34</td>
<td>600380022</td>
<td>Machine Setter</td>
<td>55L04 7</td>
</tr>
<tr>
<td>35</td>
<td>603280010</td>
<td>Grinder Operator, External Tool</td>
<td>55J63 7</td>
</tr>
<tr>
<td>36</td>
<td>603482022</td>
<td>Grinder Set-Up Operator, Surface</td>
<td>55J63 4</td>
</tr>
<tr>
<td>37</td>
<td>604380014</td>
<td>Chucking-Machine Operator, Multiple</td>
<td>55K58 6</td>
</tr>
<tr>
<td>38</td>
<td>605282014</td>
<td>Planer Set-Up Operator, Tool</td>
<td>55L31 7</td>
</tr>
<tr>
<td>39</td>
<td>605382014</td>
<td>Engraver, Tire Mold</td>
<td>55L31 5</td>
</tr>
</tbody>
</table>
40 605382026 Profiling Machine Operator 55L31
41 605682026 Tooth Cutter, Escape Wheel 55L31
42 606280014 Boring-Mill Operator, Horizontal 55H95
43 612360010 Die Setter 55H69
44 612682014 Forging-Roll Operator 55I95
45 614382010 Wire Drawer 55R41
46 615280010 Slitter Service and Setter 55R81
47 617280010 Shot-Peening Operator, Tape Control 55L02
48 621282010 A-C Check-Out Mechanic 55Q02
49 649682010 Box-Folding-Machine Operator 55F06
50 649685026 Core-Cutting and Reamer 55T66
51 649685042 Envelope-Machine Operator 55I34
52 649685114 Stitcher Operator 55F06
53 650682010 Equipment Monitor, Phototypesetting 55T13
54 651380010 Printer 55D56
55 652260010 Section Leader, Screen Printing 55T70
56 667682058 Resaw Operator 55N07
57 669680010 Machine Setter 55K97
58 669682038 Double-End-Trimmer Operator 55E98
59 673382014 Sandblaster, Stone 55N54
60 673382022 Stone Polisher, Machine 55Q25
61 673662010 Top Polisher 55Q25
62 673682030 Slab Grinder 55Q25
63 673685058 Finish-Machine Tender 55I71
64 683260018 Loom Fixer 55O62
65 683682030 Flush Weaver 55D48
66 683684010 Chain Repairer 55O56
67 683684030 Weaver, Hand Loom 55D48
68 686462010 Die-Cutting-Machine Operator 55M64
69 689260022 Section Leader and Machine Setter 55O65
70 689384014 Laboratory Tester 55Q57
71 689685026 Bouffant-Curtain-Machine Tender 55C79
72 690682062 Press Operator 55O10
73 716382018 Precision-Lens Grinder 55B70
74 721281018 Electric Motor Repairer 55Q20
75 773607018 Tester, Waste Disposal 55Q57
76 724684026 Coil Winder 55R82
77 733281010 Ballpoint Pen Cartridge Tester 55Q57
78 740684014 Decorator 55V04
79 741684026 Painter, Spray 55R68
80 754684022 Caster 55O04
81 759584010 Rubber-Goods Tester 55Q57
82 783381026 Saddle Maker 55N48
83 783684022 Leather Cutter 55S68
84 785361018 Sample Stitcher 55N51
85 786682114 Front Maker, Lockstitch 55C76
86 786682174 Lockstitch-Sewing-Machine Operator 55C76
87 805381010 Boilermaker II 55A24
88 842381010 Dry-Wall Applicator 55B05
89 850684018 Stripping-Shovel Oiler 55C06
90 853683010 Curb-Machine Operator 55B52
91 860381018 Boatbuilder, Wood 55A47
92 860381042 Carpenter, Rough 55A47
93 860681010 Carpenter II 55A47
<table>
<thead>
<tr>
<th>Code</th>
<th>Number</th>
<th>Position Description</th>
<th>Code</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>94</td>
<td>862684026</td>
<td>Plumbing Assembler-Installer</td>
<td>55C30</td>
<td>4</td>
</tr>
<tr>
<td>95</td>
<td>869684034</td>
<td>Layout Worker</td>
<td>55K30</td>
<td>6</td>
</tr>
<tr>
<td>96</td>
<td>904683010</td>
<td>Log-Truck Driver</td>
<td>55D27</td>
<td>4</td>
</tr>
<tr>
<td>97</td>
<td>914382018</td>
<td>Pumper-Gager Apprentice</td>
<td>55B75</td>
<td>7</td>
</tr>
<tr>
<td>98</td>
<td>914682010</td>
<td>Pumper</td>
<td>55T97</td>
<td>5</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


